ENGINEERING *GEOBACILLUS THERMOGLUCOSIDASIUS* FOR EFFICIENT PRODUCTION OF FUELS AND CHEMICALS

BY

JIEWEN ZHOU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor Christopher V. Rao, Chair
Professor Huimin Zhao
Associate Professor Mary L. Kraft
Associate Professor Yong-Su Jin

Abstract

The ability to grow at high temperature makes thermophiles attractive for many fermentation processes. In this thesis, we aimed to develop the thermophile *Geobacillus thermoglucosidasius* as a cell factory for fuels and chemicals production from renewable biomass. *G. thermoglucosidasius* is a facultatively anaerobic thermophilic bacterium growing between 37°C to 70°C, and can ferment diverse carbohydrates. Firstly, we engineered two strains of *G. thermoglucosidasius* 95A1 (95A1) and *G. thermoglucosidasius* C56-YS93 (C56) for improved ethanol production by employing an evolutionary engineering strategy. We eliminated lactate and formate formation in both of the strains to divert carbon source to ethanol production. However, strains were unable to grow under microaerobic conditions. Growth of 95A1 was then recovered by serial adaptation of the strain in the presence of acetate. The evolved strain of 95A1 was able to efficiently produce ethanol during growth on glucose or cellobiose. Genome sequencing identified loss-of-function mutations in adenine phosphoribosyltransferase (*aprt*) and the stage III sporulation protein AA (*spoIIIAA*). Their effect on improving ethanol production was verified by disruption of both genes. By comparison of two strains, 95A1 was a good ethanol producer and easily genetic engineered.

G. thermoglucosidasius 95A1 was furtherly developed to produce a valuable biochemical of *2R, 3R*-butandiol (*R*-BDO). Strong promoter from lactate dehydrogenase (*ldh*)_from *G. thermodenitrificans* was selected from different promoters for efficient pathway construction. The new *R*-BDO biosynthetic pathway was constructed in 95A1 through testing different combination of enzymes. As a result, an efficient pathway was obtained, which was composed of heterologous acetolactate synthase (*alsS*) from *Bacillus. subtilis* and acetolactate decarboxylase (*alsD*) from *Streptococcus thermophilus*. In order to enhance *R*-BDO production, different

fermentation conditions were optimized, including oxygen supply, temperature, inoculation time, different media and addition of yeast extract. With optimal conditions, 7.2 g/L *R*-BDO was achieved after 48 h at 55°C. Different alcohol dehydrogenase (*adh*) was also deleted separately to divert more carbon source to *R*-BDO, but it only increased the yield but not production.

Secondly, we engineered *G. thermoglucosidasius* 95A1 to secrete heterologous protein secretion since it was found to have capability of secreting proteins at high titers in our lab. To improve heterologous protein secretion, 25 signal peptides from *G. thermoglucosidasius* C56-YS93, predicted by SignalP 4.1 server, were screened and characterized for different secretory target proteins. Three thermostable hydrolase for biofuel production were selected: α-amylase (*amyE*) from *G. stearothermophilus*, endoglucanase (*eglS*) from *B. subtilis* and cellulase (*celA*) from *Caldicellulosiruptor bescii*. The optimal signal peptides for different enzymes were determined by measuring enzyme activity in supernatant of culture compared with that of their native signal peptides. One signal peptide was found to have efficient secretion with all of three enzymes.

Lastly, *G. thermoglucosidasius* 95A1 was engineered for bioethanol production by consolidated bioprocess. Efficient ethanol production from starch was achieved by transforming α-amylase secretion plasmid into the evolved strain for ethanol production. Ethanol production from cellulose was also tried in the evolved strains harboring the cellulase secretion plasmids. However, the production was inefficient, which needed more effort to be improved. In summary, the work in this thesis established this thermophile as a platform organism for fuel and chemical production from biomass.

To My parents, my husband and my son

Acknowledgements

First of all, I would like to thank my advisor Prof. Christopher V. Rao for his guidance on these projects. He not only trained my experimental skills, but also the essence to be a good researcher. He is also acknowledged for helping me to join the Department of Chemical and Biomolecular Engineering at University of Illinois at Urbana-Champaign and his group to start my academic life and his continuous help for my life. I would also like to thank my committee members, Prof. Mary L. Kraft, Prof. Huimin Zhao and Prof. Yong-Su Jin for their insightful input, constructive suggestions, and generous help.

I am also very thankful to all Rao Group members for their helpful discussions, friendship, and assistance, especially Dr. Kang Wu, Dr. Angel Rivera who worked on *Geobacillus* projects before. For the ethanol project, I would like to thank Dr. Kang Wu on her work on construction of the integration vector and Pdc plasmid, and optimization of fermentation conditions. She is also appreciated for her work of developing thermostable fluorescence protein, which was used for my promoter characterization in *R*-BDO project. For the protein secretion project, I would like to thank Dr. Angel Rivera, who worked on signal peptides prediction and identification of total protein secretion on different substrates, and Dr. Hana Rao, James Orr for their help and suggestion on the project. For fermentation experiments, I would like to thank Dr. Shuyan Zhang for her help in setting up bioreactors. Dr. Ahmet Badur is acknowledged for his help in protein purification (not included in the thesis). Dr. Kori Dunn and Xiaoyi Wang are both appreciated for help me ordering different enzymes, chemicals and supplies. I also thank my other Rao lab members: Dr. Charles Rutter, Dr. Sujit Jagtap, Dr. Ehab Ammar, Yuki Kimura, Khushnuma Koita, Santosh Koirala, Payman Tohidifar, Geethika Yalamanchili, Sichong Pei, Anshu Deewan and Girija Bodhankar.

I would like to thank Prof. Isaac Cann from our campus to share the genomic DNA of *Caldicellulosiruptor bescii* for cellulase gene cloning. I would like to thank Dr. Alvaro Hernandez from Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign for his help in next generation sequencing.

Finally, I would like to express my most sincere gratitude to my parents, my husband and my son for their constant support and unconditional love. Love all of you forever!

Table of Contents

Chapter 1 Introduction

1.1 The genus *Geobacillus*

The genus *Geobacillus* is a group of thermophilic, Gram-positive, facultatively anaerobic, spore-forming *bacilli* [1]. *Geobacilli* grow at a wide range of temperatures from 37°C to 80°C and grow optimally between 50°C and 60°C. They are readily isolated from high-temperature environments including hot springs [2], compost [3,4] , oil fields and volcanoes [5]. However, they also distribute in a wide range of moderate- and low-temperature environments including soil [6,7] and permanently cold ocean sediments [8]. They are capable of utilizing various carbon sources including hexose and pentose sugars and oligomers[9]. The genus was introduced by Nazina et al. in 2001[5] through reclassifying a separate subgroup (group 5) of thermophilic *bacillus* strains, according to results of a thorough and polyphasic examination. There are 15 validly described species within the genus of *Gebacillus*. Aliyu et al. [10] revaluated the phylogenetic relatedness among the 63 *Geobacillus* strains by using whole genome approaches. They demonstrated that the genus of *Geobacillus* should be divided into two clades because of different nucleotide base composition, and proposed clade II as a new genus of *Parageobacillus*.

1.1.1 Application of *Geobacillus*

Thermophilic microorganisms attract more and more interest on industrial production because of advantages of their high temperature growth: 1) the possibility of contamination in bioprocess is minimized; 2) the cost of cooling in process is reduced; 3) the separation of product will be easier if it's volatile [11]. Moreover, *Geobacillus* are capable of utilizing various carbon sources and growing at high rate and density [1].

Based on these advantages, *Geobacillus* species have significant biotechnological applications. Firstly, they are new ideal producers for biofuel and important chemicals. Genetic tools are developed to engineering *Geobacillus* for improved production of biofuel, such as ethanol [12], isobutanol [13]. They are also used for production of important chemicals, such as lactic acid [14], exopolysaccharides [15], 2,3-butanediol [16]. Secondly, they are provider of important thermostable enzymes: amylase [17,18], lipase [19,20], protease [21,22], carboxylesterases [23,24], DNA polymerase [25], reverse transcriptase [26]. Besides, some of them can be involved in biodegradation [27,28,29,30], biorefinement [9], biomaterial [14,31] and biocontrol [9].

1.1.2 Genetic tools for *Geobacillus*

The biotechnological potential of *Geobacillus* has pushed the development of reliable genetic tools for engineering metabolic pathways to increase the production, or for better understanding of their properties. Due to the relatively close phylogenetic relation to the genus of *Bacillus*, many tools used in *Geobacillus* are from or modified from methods for *Bacillus*. In the meanwhile, novel tools for *Geobacillus* are also developed because some from mesophilic *Bacillus* are not thermostable or compatible. Continues improving on genetic engineering of these thermophilic bacteria was overviewed [1,9,32], which included transformation, vectors, antibiotics, gene integration and knockout.

1.1.2.1 Gene transfer

In order to genetic engineer a strain, an efficient method of gene transfer is essential. *Geobacillus*, as thermophilic Gram-positive bacteria, is recalcitrant to genetic transformation, because the thick peptidoglycan layer and the low permeability of the plasma membrane hinder

DNA transferring to the cells [1]. The main methods developed for *Geobacillus* are protoplast transformation, electroporation and conjugation.

Protoplast transformation is the earliest successful transformation method reported in *Geobacillus* by Imanaka et al in 1982 [33]. Their protocol was modified from the one applied for *B. subtilis*. It includes three steps: 1) protoplast preparations by treating cells with lysozyme; 2) polyethylene glycol (PEG) treatment to induce DNA transferred into protoplast; 3) protoplast regeneration on plates with antibiotic. This method is considered to be time-consuming and laborious, because the fragile protoplast needs to be prepared freshly before every transformation. In the genius of *Geobacillus*, it was only reported to apply in *G. stearothermophilus* [33,34,35].

The application of electroporation in *Geobacillus* began in early 1990s [36,37] which is widely used in different bacteria. This method contains three steps: 1) preparation of electro-competent cells. Cells are often harvested during exponential phase and then washed with very low ionic strength buffer [38]. 2) DNA transfer by electroporation. Brief exposure of cells with DNA to strong electric field makes formation of tiny holes in the cell membrane, which DNA can pass through [39]. 3) Cell recovery. Cells need to recover in appropriate liquid medium before growing on selective plates. To maximize the transformation efficiency, different conditions can be optimized, including growth time, electroporation buffer, electric field, DNA concentration and recovery time [37,38]. The reports about electroporation conditions optimization are rare. The highest transformation efficiency reported in *Geobacillus* is $2.8 * 10^6$ transformants per microgram DNA in a strain of *G. thermodenitrificans* [40]. It is consider to be a convenient and efficient transformation technique, which has applied in *G. stearothermophilus* [41], *G. thermoglucosidasius* [12,42,43] and *G. thermodenitrificans* [36,37].

Another effective method of gene transfer in *Geobacillus* is conjugation. It often needs a donor strain of *Escherichia coli*, a plasmid with an origin of transfer, and a transfer gene [44]. The method consists of four steps: 1) transformation of mobile plasmid with desired gene to the donor cell of *E.coli*; 2) culture of donor cells and recipient cells separately; 3) conjugative transfer of the plasmid from the donor to the recipient by co-culturing them at 37°C; 4) selection of recipient cells with desire plasmid on selective plates at its appropriate temperature. This method is difficult and time-consuming, but it is very effective for some strains. *G. kaustophilus* HTA426 [45] and *G. thermoglucosidasius* C56-YS93 (in our study, published), which are recalcitrant to electroporation, can be transferred through conjugation.

Another potential method for *Geobacillus* is natural transformation, which utilizes the ability of bacteria to uptake naked DNA from environment [46]. Many bacterial organisms have natural competence. It is found in thermophilic bacteria, such as *Streptococcus thermophilus* [47], *Thermus aquaticus*, *Thermus thermophilus* [48], and *Thermoanaerobacterium saccharolyticum* [49]. *B. subtilis* is one of the best studied model organisms for research on natural transformation in Gram-positive bacteria. Other *Bacillus* species, *B. amyloliquefaciens*, *B. licheniformis* [46] and *B. cereus* [50] also exist natural competence. Natural competence is controlled by the competent genes, which are also found in *Geobacillus* species [1]. Therefore, natural competent transformation is another possible method of gene transfer in *Geobacillus*.

1.1.2.2 Cloning vectors for *Geobacillus*

The development of cloning vectors, which can replicate and maintain in multiple generations, is also necessary for genetic engineering. The cloning vector of *Bacillus*, pUB110, is one of the earliest plasmids transformed to *Geobacillus* [33,51]. It can replicate in *Geobacillus*,

but only when the temperature is below 55°C [9]. To find out origins of replication working at high temperature, the best way is sequencing the plasmids in different *Geobacillus*. The sequenced plasmids are limited, including pSTK1, pGS18, pGTG5, pGTD7, pBST1 and pTB19 [35,41,52,53], but the number should increase as more and more *Geobacillus* bacteria are genome sequenced. Shuttle cloning vectors were constructed by combining those origins of replication with *E.coli* cloning vectors. The *G. stearothermophilus* shuttle vector pBST22 contains the replication origin of pBST1 and a thermostable kanamycin resistance gene (TK101) in *E.coli* cloning vector of pUC19. This vector is kanamycin resistant in *Geobacillus* while ampicillin resistant in *E.coli*. It is stable and selective up to 70°C but it lacks the multiple cloning site and blue-white selection [35] [54]. Another vector named pUCG18 was then constructed, which has the same origin of replication and resistance gene from pBST22. It also incorporated the multiple cloning sites and β-galactosidase gene from pUC18 [42]. Its derivative, pUCG3.8, was constructed later to increase transformation efficiency through reducing the size [55]. Plasmid of pNW33N is another widely used shuttle vector in *Geobacillus* for protein expression. It harbors the replicon of pBC1 from thermophilic *B. coagulans* and chloramphenicol resistance gene from *Staphylococcus aureus* plasmid pC194, which is only stable below 60°C [40]. The moderate thermophilic replicon of pUB110 is often used for temperature sensitive suicide vector for gene integration [12].

1.1.2.3 Selection systems

Selection systems are required for identification of plasmid transformation or gene knockout. Antibiotic selections are the most commonly used ones combined with plasmids carrying related selection markers. There are a lot of antibiotics used in bacteria, but only a few of them can be used in thermophilic bacteria. This is limited by the requirement of thermostability for both of

the antibiotic and the resistance marker. Kanamycin is the most popular and thermostable antibiotic [1,12,42,55]. The thermostable kanamycin marker (TK101) was obtain by spontaneous mutation of mesophilic marker at 63°C [35]. It functions up to 70°C, and works in both of *Geobacillus* and *E.coli*. Chloramphenicol is another popular antibiotic used in *Geobacillus* [12,13,40]. As mentioned previously, pNW33N carries the resistance marker, which functions below 60°C. Tetracycline can also be applied as selection agent for the transformation of *Geobacillus* up to 65°C [33].

Antibiotic-free selection systems are developed simultaneously because of the preference in industry. This kind of systems is often based on the complementation of auxotrophy. Auxotrophic selectable markers are the essential genes in a particular metabolic pathway, which are inactivated in host strains. The most widely used is the *pyrE* or *pyrF* marker, which are genes encoding orotate phosphoribosyltransferase and orotidine 5'-phosphate decarboxylase respectively, involving in de novo biosynthesis of pyrimidine-related metabolites (UMP, UDP and UTP) [11,56]. The disruption of either of them will result in the lack of those essential metabolites, which can be overcome by supplementation of uracil. Uracil can be converted into UMP by uracil phosphoribosyltransferase. In the meanwhile, PyrE and PyrF can catalyze the conversion from 5-fluoroorotic acid (5-FOA) to 5-fluorodeoxyuridine 5'-monophosphate, which is toxic to cells. Therefore, $pyrE^-$ or $pyrF^-$ causes uracil auxotrophy and 5-FOA resistance, which is applied as a counter selection system with a *pyrE* or *pyrF* marker on plasmids. This system has been demonstrated in many microbes for marker-free gene knockout, and it's also applicable in *G. kaustophilus* HTA426 [11] and *G. thermoglucosidasius* NCIMB 11955 [56]. Another counter-selection system was described recently in *G. thermoglucosidasius* by using β-glucosidase (Bgl) and the synthetic substrate X-Glu (5-bromo-4-chloro-3-indolyl-β-d-glucopyranoside) [57]. Its

principal is that the Bgl cleavage products of X-Glu are toxic. Although the wild type has native *bgl* gene on the chromosome, it can still be distinguished from the one with overexpressed Bgl through increasing the concentration of X-Glu.

1.1.3 *G. thermoglucosidasius*

G. thermoglucosidasius is a species within clade II of the genus *Geobacillus* [10], so it has a different name of *Parageobacillus thermoglucosidasius*. It is a facultative anaerobe, which can grow at a wide range of temperature from 37 to 68°C and pH from 6.0 to 8.0. It utilizes various carbon sources including hexose, pentose sugars and short-chain oligosaccharides, such as gluco-oligosaccharides and manno-oligosaccharides. Most of strains can also hydrolyze gelatin, pullulan, and starch.

The genetic tools for *G. thermoglucosidasius* are well developed. The most frequently used method of transformation is electroporation [12,42,43], but the efficiency is low for some strains, such as *G. thermoglucosidasius* C56-YS93 [43]. To engineer its pathways, Cripps *et al.* [12] developed a technique of marker-free gene deletion in G. *thermoglucosidasius* by utilizing homologous recombination (Figure 1.1). Integration vector was constructed firstly for gene knockout, which contains two DNA fragments amplified respectively from upstream and downstream of the target gene. The whole vector was integrated into the genome through single crossover and marker-free mutants were obtained by double, reciprocal crossover recombination. This method is laborious because marker-free mutants need to be screened from wild type by PCR. Recently, an improved method was reported, which incorporated counter-selection using β-glucosidase and the synthetic substrate X-Glu [57]. It prevented the re-creation of wild type and furtherly avoid screening by PCR. Until now, four strains of *G. thermoglucosidasius* have

completed genome sequences online: *G. thermoglucosidasius* C56-YS93, *G. thermoglucosidasius* DSM 2542, *G. thermoglucosidasius* TNO-09.020, *and G. thermoglucosidasius* Y4.1MC1. Based on the above, this species is engineered for biotechnological application.

1.2 Engineering production of biofuel and chemical

1.2.1 Biofuel

The depletion of fossil fuels and climate change has caused a high demand for energy security by generating energy from sustainable resources. Among different strategies, biological conversion of plant-derived lignocellulosic materials into biofuels (cellulosic biofuels) is considered as one of the most promising routes to solve these problems [58]. First of all, the raw material, cellulose biomass, is sustainably available at low cost and the production of cellulosic biofuels was found to reduce the greenhouse gas emission level. In addition, it was proposed that cellulosic biofuel industry had no negative effect on food supplies [59].

Lignocellulosic biomass is mainly composed of three polymers: cellulose, hemicellulose, and lignin, which needs pretreatment and scarification before conversion into biofuels (fermentation). The pretreatment can be a chemical pretreatment step involving acid or alkali, a physical step involving high pressure or grinding, or a combination of both. It makes lignocellulosic material amenable to scarification. Scarification is a process of hydrolyzing those polymers into simple sugars. This process can be finished by chemical method or biological method [60].

Different with that of starch and sugarcane, a mixture of sugars including hexose, pentose and oligomers are released from lignocellulosic biomass after pretreatment [61]. To make the cellulosic biofuel industry more economically competitive, the producing host should be able to

consume all sugar components. Currently, *Saccharomyces cerevisiae*, also known as baker's yeast, is the most commonly used organism for cellulosic biofuel production [62]. Unfortunately, wild type *S. cerevisiae* strains cannot utilize pentose sugars and oligomers, such as xylose and cellobiose, the abundant carbohydrate components in lignocellulosic biomass. Heterologous pathways have to be introduced to enable various sugar metabolisms. Although steady progress has been made after intensive engineering work, the consumption rate is still low, compared with that of glucose [63,64]. In addition, many attempts have been made to engineer mesophilic bacteria, such as *Escherichia coli* [65], *Zymomonas mobilis* [66], to produce biofuel. However, these organisms grow at moderate temperature (30-37°C), while pretreatment of cellulosic biomass generally occurs at temperatures higher than 55°C [67]. Therefore, a cooling step is necessary to enable efficient biofuel fermentation, leading to increased cost and lower process efficiency [68].

1.2.2 Engineering for ethanol production

Increasing efforts have been made to use thermophilic microorganisms as the biofuel producers, especially ethanol, which is the most popular and well-established biofuel. High-temperature fermentations by thermophiles have potential advantages for the ethanol production from lignocellulosic biomass. In addition, some thermophilic microbes are capable of ferment cellulosic biomass rapidly without adding enzymes [69], which helps carry out consolidated bioprocessing (CBP). CBP combines biomass degradation and sugar fermentation in a single step without enzyme addition, is widely considered to be the ultimate low-cost configuration for cellulosic biofuel industry [70]. Engineered thermophilic organisms, used for ethanol production, include obligate anaerobic bacteria, *Thermoanaerobacterium saccharolyticum*, *Thermoanaerobacter ethanolicus*, *Thermoanaerobacter mathranii*, *Clostridium thermocellum*

and *Caldicellulosiruptor bescii*, facultatively anaerobe *G. thermoglucosidasius*, the anaerobic archaeon *Pyrococcus furiosus* and the thermotolerant yeast *Ogataea polymorpha* [69].

G. thermoglucosidasius is a promising bioethanol producer, because it is a natural thermophilic ethanologen and it can ferment different sugars. During sugar fermentation, it produced a mixture of lactate, acetate, formate, and ethanol during sugar fermentation in oxygen limited conditions (Figure 1.2) [12,71]. To improve ethanol production in *G. thermoglucosidasius*, Cripps *et al.* [12] deleted the genes encoding lactate dehydrogenase and pyruvate formate lyase and expressed the gene encoding pyruvate dehydrogenase by using the oxygen-insensitive promoter for lactate dehydrogenase from *G. stearothermophilus*. This design eliminated mixed-acid fermentation and resulted in ethanol being produced at high yields. Another strategy to improve ethanol was introduction a new pathway by overexpression of pyruvate decarboxylase from *Zymomonas mobilis* to convert pyruvate to acetaldehyde. Interestingly, although *Z. mobilis* cannot grow at high temperature, its *pdc* gene product could retain the enzymatic activity up to 52^{o}C [72].

1.2.3 Engineering for isobutanol production

As fossil fuel alternatives, isobutanol has more advantages over ethanol, such as higher energy density, lower oxygen content, lower Reid Vapor Pressure, and low water solubility [9]. *G. thermoglucosidasius* was then metabolic engineered for isobutanol production by Lin *et al*.[13]. In this study, thermostable enzymes required for isobutanol biosynthesis were screened and characterized. The production of 3.3 g/L isobutanol from glucose at 50^{o}C was achieved by overexpressing of *B. subtilis* acetolactate synthase, *G. thermoglucosidasius* ketol-acid

reductoisomerase and *Lactococcus lactis* ketoisovalerate decarboxylase. *G. thermoglucosidasius* is proved to be a platform organism for different biofuels and chemicals.

1.2.4 Engineering for 2,3-butanediol production

2,3-Butanediol (BDO) has attracted growing attention due to its important industrial applications [73,74]. As reported, approximately 32 million tons of the BDO derivatives per year are potentially needed for the market, which equal about $43 billion. It is used as an antifreeze agent because of its low freezing point. It is also an important chemical intermediate used for the production of synthetic rubber, fuel and drugs [73]. BDO contains two stereo centers and has three isomeric forms in nature, (*2S,3S*)-butanediol (*S*-BDO), meso-butanediol (*meso*-BDO), and (*2R,3R*)-butanediol (*R*-BDO). A lot of microorganisms can natively produce BDO, especially *Klebsiella* and *Enterobacter* species, which can accumulate BDO to high titers. However, these native producers are not commercial friendly enough, since they are pathogenic or the BDO synthesized is generally a mixture of stereoisomers [75]. Therefore, some industrially friendly hosts are being developed for BDO production, such as *E. coli* [76], *B. subtilis* [77] and *S. cerevisiae* [78,79].

Enantiopure building blocks are essential for synthesis of high-value compound used in pharmacy or agriculture. *R*-BDO is paid attention because of its specific application, e.g. the use of *R*-BDO as an antifreeze agent and a low freezing point fuel [80]. Compared to chemical synthesis, the microbial production of *R*-BDO is more effective, which has been developed in different hosts. The first report is in 2009, Yan *et al.* [81] introduced a synthetic pathway of *R*-BDO in the engineered *E. coli*. It produced 6.1 g/L of *R*-BDO with a high enantiopurity; the production in *E. coli* was significantly improved to 115 g/L by fed-batch culturing [76]; High

level of enantiopure *R*-BDO is also achieved in metabolic engineered *Saccharomyces cerevisiae* [82,83]; *Paenibacillus polymyxa* was capable of producing high purity *R*-BDO to 36.92 g/L by using raw inulin extract as carbon source after medium optimization [84]. Most of the reported producers are mesophilic with the optimal growth temperature between 30°C to 40°C. Because of the benefits of high temperature fermentation, thermophiles are explored for BDO production. Some of *B. licheniformis* and *B. subtilis* strains can produce BDO at 50°C [85,86,87], but only *B. licheniformis* BL1 was reported to produce high enantiopure *R*-BDO. It was also the only reported strain engineered for improved *R*-BDO production at high temperature. The wild type was actually a lactic acid producer, which became an *R*-BDO producer (13.8 g/L) after blocking the lactate synthetic pathway [88]. Reports about BDO production in *Geobacillus* are rare. The only one is *Geobacillus* sp. XT15, with a natural ability of BDO exporting [16].

According to genomic analysis, *G. thermoglucosidasius* only has the pathway to produce *R*-BDO (Figure 1.2). The pathways include four reactions, one of which is a spontaneous step requiring O_2. The other three reactions are catalyzed by two enzymes, acetolactate synthase (*alsS*) and (*2R,3R*)-butanediol dehydrogenase (*bdhA*). AlsS (EC 4.1.3.18) catalyzes the first reaction from pyruvate to 2-acetolactate. In a previous study [13], the native AlsS in *G. thermoglucosidasius*, together with other two putative enzymes from different thermophiles, was compared with the one from *B. subtilis*. AlsS from *B. subtilis* had been proved in different studies to have high specific activity on pyruvate for acetolactate formation without end-product inhibiting [89,90]. That study furtherly confirmed that *B. subtilis* AlsS had high specific activity up to 60°C. However, the native AlsS in *G. thermoglucosidasius* and others had little activity on pyruvate. The second reaction is spontaneous, which requires oxygen to covert acetolactate into diacetyl. The last two reactions are catalyzed by a bifunctional enzyme (*2R,3R*)-butanediol

dehydrogenase (EC 1.1.1.4). In other hosts, such as *B. subtilis*, there is an alternative way between acetolactate and (*R*)-2-acetoin, which is catalyzed by acetolactate decarboxylase (AlsD, EC 4.1.1.5) without the requirement of O_2. *B. subtilis* AlsD is often introduced to construct a new pathway for BDO production for mesophilic microorganism [82].

1.3 Protein secretion

Secretion is the process of exporting proteins to the cell exterior, which happens in all living organisms. In bacteria, 5–10% of the proteins encoded on their chromosomes were secreted [91], which are involved in breakdown of polymeric substrates, cell-wall synthesis, cell division, cell-to-cell communication, detoxification of the environment, and killing of potential competitors [92,93]. In Gram-negative bacteria, such as *E. coli*, there are evolved specialized substrate-specific protein secretion pathways (e.g. types I, III and IV), while in Gram-positive bacteria, there are no specialized secretion pathways but the ubiquitous Sec-dependent (Sec) and twin-arginine translocation (Tat) pathways for the protein secretion [92].

Secreted protein has several advantages compared to intracellular protein: its purification is much more simplified; the structural authenticity improves; the possibility of contamination decreases. Therefore, candidates with good capacity of protein secretion are developed and engineered for extracellular production of heterologous proteins. *B. subtilis* is widely used in industry for enzymes production, which can secrete proteins at high titers (20-25 g/L) [9]. So far, there are four distinct pathways for protein secretion: Sec-SRP cooperation pathway, Twin-arginine translocation (Tat) pathway, ATP-binding cassette (ABC) transporters and a pseudopilin export pathway for competence development [94]. Protein secretion in *B. subtilis* is well studied and it has close relation to *G. thermoglucosidasius*, so it is an appropriate reference for study of protein secretion in *G. thermoglucosidasius*.

1.3.1 Secretion pathways

1.3.1.1 Sec- SRP pathway

Most of secreted proteins are transported across the cytoplasmic membrane via Sec pathway. The main secretion pathway in *B. subtilis* consists of SRP-protein targeting system to the cell membrane and the Sec protein translocation machinery across the cytoplasmic membrane [94]. It can be divided into three functional stages (Figure 1.3): targeting, translocation, folding and release.

Targeting is a crucial step in Sec pathway, taking place right after precursor protein is synthesized. The precursor protein has an N-terminal signal peptide containing 20-30 amino acids, which is recognized by Signal recognition particle (SRP). SRP is a highly conserved and essential RNA–protein complex that interacts with hydrophobic regions of signal peptides. The composition of the SRP varies in different organisms. In the genus of *Bacillus*, it consists of an RNA molecule (small cytoplasmic or scRNA) as the backbone and two attached proteins: Ffh, similar to the 54kDa protein of eukaryotic SRPs, and HBsu, a histone-like DNA-binding protein [92,94].

The precursor protein is passed through the secretory translocase, after the SRP delivers it to a membrane-bound docking protein (FtsY). The N-domain of the signal peptide interacts with phospholipids in the membrane because of different charges, which leads looping insertion of the H-domain [94]. The first part of the precursor is then pulled through the translocator after the H domain unloops [95]. The Sec translocases is composed of SecA, SecYEG and SecDF. SecA is a peripheral membrane-associated ATPase and functions as the “motor” to provide energy for the precursor translocation. It also helps target the precursor-SRP complex to the translocator, the

integral membrane pore, formed by integral membrane proteins of SecY, SecE, SecG and SecDF [96].

During or shortly after translocation, the mature protein is released via cleavage of signal peptide by SPases. The secreted protein is folded or degraded under the quality control by extracellular chaperones, such as HtrA, HtrB and WprA.

1.3.1.2 Twin-arginine translocation (Tat) pathway

Tat pathway is distinguished from Sec-SRP pathway: 1) Signal peptides of Tat pathway have specific RR-motif for recognition; 2) Sec pathway secretes unfolded proteins across the cytoplasmic membrane, while Tat pathway is capable of transporting folded proteins or even multimeric enzyme complexes [97]. This pathway is named because of twin arginine residues on N terminal of signal peptides, which are required for targeting preprotein to Tat pathway [98]. It functions as a channel for secretion of Sec-incompatible proteins, which fold too fast or tightly [99]. In *E. coli*, it transports certain periplasmic precursor proteins that need cofactors in the cytoplasm, such as flavins, molybdopterins and iron–sulfur clusters [100]. The Tat pathway is found in many eubacteria, some archaea, and a few plant mitochondria, but their cellular components are different [97]. In Gram-negative bacteria, the main components of the Tat translocation machinery are TatA, TatB and TatC, while Gram-positive bacteria only have TatA and TatC, such as *B. subtilis* [101,102]. TatA and TatC are forming a docking complex, which interacts with the protein right before translocation happens [103].

1.3.1.3 ATP-binding cassette (ABC) transporters

ATP-binding cassette (ABC) transporters may export some of secretory proteins or antimicrobial peptides, which are incompatible with Sec or Tat pathway. They are found in eukaryotes and prokaryotes and are composed of a large superfamily of multi-subunit permeases.

They are usually involving in import or export of various molecules through membranes, such as ions, amino acids, peptides, antibiotics, polysaccharides, proteins. [94]

1.3.2 Signal peptides

Signal peptide (SP) is a regulatory element at the N-terminus of secretory protein. It is essential for secretory proteins targeted to Sec or Tat pathway. Different signal peptides don't have conserved amino acids, but they still share some common characteristic features. Most SPs consist of three distinct regions: the positively charged N domain, the hydrophobic core region of H domain and the hydrophilic signal peptidase (SPase) recognition site of the C domain [104]. The N-domain contains one or two positively charged residues (such as arginine or lysine) and is involved in the translocation machinery. It interacts with negatively charged phospholipids in the membrane, so it is thought to determine the final orientation of SPs in membrane [93,105]. The H-domain, following the N-domain, is hydrophobic core of SP. It is proposed that it can form an α-helical structure in the membrane [106] and unlooping of the structure makes the whole SP insert into the membrane [107]. The C-domain, following the H-domain, contains the specific signal peptidases cleavage site where the signal peptide will be removed from the preprotein during or shortly after translocation. There are two types of signal peptidases. Type I signal peptidases cleave most of secretory proteins and the relative C-domain ends with Alanine-X-Alanine (AXA) at the −3 and −1 position relative to the cleavage site [92]. In *B. subtilis*, there are five enzymes, SipS, SipT, SipU, SipV and SipW. Type II signal peptidases only cleave signal peptides from lipoproteins. In *B. subtilis*, there is only one type II SPase of LspA [105].

SPs are summarized to possess three functions [105]: 1) inhibit the preprotein folding in cytoplasm to maintain translocation competence and avoid activating secretory enzyme in

advance; 2) interact with other secretion machinery components and direct the translocation; 3) serve as a topological determinant for proenzymes in the membrane.

1.3.3 Extracellular proteases

High level of proteases are exported into cell wall or the surroundings, which can be divided into two groups according to their functions: One consists of quality control proteases playing a key role in cell homeostasis; the other consists of feeding proteases digesting proteins and peptides into small molecules as nutrients [108]. The former ones monitor the secreted proteins from the cells and degrade misfolded ones. In *B. subtilis*, the main quality control proteases are the serine proteases WprA, HtrA and HtrB [109,110], which are the main barrier to the production of heterologous protein secretion [92].

Membrane-bound HtrA and HtrB in *B. subtilis* are regulated by a two-component system of CssR and CssS (CssRS: Control of secretion stress Regulator and Sensor). The system can be stimulated by secretion stress and heat shock, which will then activate the two proteases [111]. By comparing the mutants of these genes, it shows HtrA in *B. subtilis* may play a second role of helping recover the configuration of misfolded proteins, while targeting them for degradation [112]. Besides, HtrA and HtrB are crucial for the cell integrity under non-stress conditions [113]. Wall-associated protein A (WprA) is a cell-wall-bound quality control protease. It is proteolytically processed into two products: the one from the N-terminal portion is a serine protease involved in the degradation of non-native secretory proteins; the other from the C-terminus is like a propeptide involved in the folding and control of its activity [110]. In *B. subtilis*, there are seven feeding proteases: NprB, AprE, Epr, Bpr, NprE, Mpr and VprA. All are involved in proteolytic degradation of native or heterologous extracellular proteins.

1.3.4 Heterologous protein secretion in *Geobacillus*

1.3.4.1 Bacterial hosts for heterologous proteins secretion

Efficient protein secretion has important biotechnological applications for enzyme production because of its commercial advantages. Therefore, promising host organisms for the secretory production of heterologous proteins are developed and optimized. Due to the lack of an outer membrane, Gram-positive bacteria are considered as good candidates for protein secretion. Members of the *Bacillus* genus have been investigated intensively for industrial proteins production of α-amylases, proteases, lipases, and other macromolecular hydrolases, which have capacity of naturally secreting a wide range of hydrolytic enzymes at high concentration [114]. *B. subtilis* is paid more attention not only because of protein secretion ability but also it is a GRAS organism (generally recognized as safe) [94]. An α-amylase from *B. licheniformis* was secretory expressed in *B. subtilis* with extracellular activity of 2012 U/ ml [115]. Exoproduction of an isomalto-dextranase from an *Arthrobacter sp.* in *B. subtilis* reached approximately 4.6 g/ l in culture broth [116]. Attention has also been paid to other *Bacillus*, such as *B. brevis* [117,118], *B. licheniformis* [119] and *B. megaterium* [120]. *L. lactis* is another Gram-positive organism considered as an attractive host for the secretion biologically useful proteins, which is a model food-grade lactic acid bacterium extensively used in dairy fermentations [121,122]. Compared to *B. subtilis*, the main advantage of *L. lactis* is that only one housekeeping protease exists in laboratory strains [123]. Other bacteria are used for heterologous protein secretion, including *Streptomyces lividans* [124], *Corynebacterium glutamicum* [125], some species in *Lactobacilli* [126], *Bifidobacterium* and *Clostridium* [124].

1.3.4.2 Strategies to increase secretion of heterologous proteins

Signal peptide directs the targeting of secretory protein to the translocation machinery and interacts with intracellular chaperones such as SecB, Ffh or CsaA will bind and direct to the translocation machinery. Therefore, it plays a pivotal role in heterologous protein secretion, which is always optimized for improvement. The strategy is to construct SPs library and screen for the optimal ones. The library is composed of different SPs from different strains [119,127,128] or SPs with different mutations [129,130,131]. The method is very effective to increase secretion efficiency for one protein. However, according to extensive study in *B. subtilis*, the optimal signal peptides for different proteins are always unrelated. In other words, SPs need to be rescreened for different protein targets.

Another effective strategy is deleting extracellular quality control protease and feeding protease, since they are the largest barrier for heterologous protein secretion. Wong and colleagues [132] constructed multiple-protease deficient strains, one of which was deleted the seven feeding protease and WprA. They found it was effective to improve the production of single-chain antibodies [132]. In the further study of ten-protease deficient strains [113,133], it was found: the extracellular proteases also degraded many native secreted proteins, lipoproteins and membrane proteins; quality control factors of PrsA, HtrA, and HtrB were indeed substrates of multiple extracytoplasmic proteases; the improvement of heterologous protein in protease deficient strains was not a simple result of reduced protein degrading, but also related to reduced proteolysis and improved posttranslocational protein folding. Although the relation of these factors is still not completely clear, the deletion of multiple proteases is still an effective way to increase secretion efficiency.

Since protein secretion is based on protein expression, the ways to improve protein expression are also applicable, such as promoters, gene copy numbers. Besides, increasing levels of endogenous molecular chaperones can also enhance proteins secretion.

1.3.4.3 *Geobacillus* for heterologous proteins secretion

Thermostable proteins secreted from *Geobacillus* are drawn attention because of their biotechnological application, such as amylase [17,18], lipase [19,20] and protease [21,22]. Most of these enzymes are heterologously expressed in mesophilic hosts, such as *E. coli*. These hosts are well developed for protein expression, but still have drawbacks to express proteins from thermophiles. Some protein cannot fold correctly in these mesophilic hosts possibly because of large temperature difference and distant phylogenetic relationship [134]. Moreover, protein expression is effected by codon usage bias and lacking of co-factors [1]. From another point of view, members of *Geobacillus* are potential candidates for protein secretion, because it is a closed relative of *Bacillus*. By analysis of genomes and comparison with *B. subtilis*, it is found members of *Geobacillus* have the main components of Sec and Tat pathways for exporting proteins [55,135]. Therefore, *Geobacillus* attracts great interest for protein secretion. The secretion of native α-amylase and xylanase in *G. thermoleovorans* was improved by optimization of culture conditions [136,137,138,139]. A truncated cellulase from *Pyrococcus horikoshii* and a α-amylase from *G. stearothermophilus* with their native signal peptides were secreted successfully in *G. kaustophilus* HTA426 [134]. A module system for inducible secretion of endoglucanase *cel5A* from *Thermotoga maritima* was constructed in *G. thermoglucosidasius* NCIMB 11955, which allowed interchange of promoters, signal peptides and genes [55]. Some *Geobacillus* species could export xylanase for degrading xylan [43]. Recently, secretion of

heterologous xylanase from *G. thermodenitrificans* with its own signal peptides was achieved in *G. thermoglucosidasius* [43,135].

1.3.5 Heterologous protein secretion of α-amylase

α-Amylase (EC 3.2.1.1, 1,4-α-D-glucan glucanohydrolase) is an important enzyme for starch hydrolysis in the brewing and sugar industries. It cleaves randomly internal α-1,4-glucosidic bonds of glycogen, starch and related polysaccharides and produces reducing sugar [140]. α-Amylase has a wide range of sources from bacteria, fungi, plants and animals, while the bacterial and fungal ones are used in industry. In bacteria, *Bacillus* sp. is the most widely used one for thermostable α -amylase, such as *B. subtilis*, *B. licheniformis* and *B. amyloliquefaciens* [141]. Another thermostable one for industrial applicaiton is from *G. stearothermophilus*, which previously belonged to *Bacillus* with a name of *B. stearothermophilus* [142].

α-Amylase from different sources were expressed, characterized and optimized for commercial purpose [139,143]. Besides, the *amyE* gene is also used as a reporter gene, because it is easily detected.

1.3.6 Heterologous protein secretion of cellulase

Cellulose is the primary component in lignocellulosic biomass. It is a polysaccharide consisting of a linear chain of D-glucose units connected with β-1,4 glycosidic bonds. These linear chains were tightly packed through hydrogen bonding within and between neighbor chains, which forms a rigid crystalline structure [144]. Cellulases are glycoside hydrolase responsible for the degradation of cellulose through hydrolyzing β-1,4-glucosidic bonds between glycosyl residues. There are three different types of cellulase, which are working synergistically: 1) endoglucanase (endo-1,4-β-glucanase), randomly cleaves the internal bonds of the chain, 2) exoglucanase (exo-1,4-β-glucanase), attacks the reducing or non-reducing end of the exposed

chain, and 3) cellobiose (β-glucosidase), converts cellobiose to glucose [144]. In many cellulolytic microorganisms these enzymes occur in two alternative states: some are expressed in high amounts and exported to extracellular environment in a free style; others form a multicellulase complex, known as cellulosome, associated with the cell surface.

Conversion of cellulosic biomass to fermentable sugars is the major bottleneck in cellulosic biofuel production. High cost of cellulase used for enzymatic hydrolysis of cellulose is still a cost-limiting factor for biomass conversion [145]. Cellulase used in industry is mostly from filamentous fungi, which is not thermostable. However, thermostable cellulase attracts more and more interests because high temperature enzymatic hydrolysis offers many advantages: higher specific activity, higher stability, lower risk of microbial contamination, higher mass transfer rate because of lower fluid viscosity, and greater flexibility in the bioprocess [146]. Thermostable cellulase exists in some of thermophilic fungi or thermophilic bacteria. In general, cellulases from bacteria have more advantages than those from fungi: higher activity, less feedback inhibition, better stability under harsh conditions [147].

The thermophilic cellulolytic bacterium *Clostridium thermocellum* is extensively studies for its elaborate cellulosome complex and the separate free cellulase system, including endoglucanase and exoglucanase. Its cellulases are well-known for their increased thermostability, high efficiecny and substrate specificity, which have been heterologously expressed and secreted in different hosts [148,149,150]. Thermostable cellulases can also be found in several thermophilic bacteria, belonging to the genera *Bacillus*, *Geobacillus*, *Caldibacillus*, *Acidothermus*, *Caldocellum*, and *Clostridium*, which are summarized in reviews [147,151,152,153].

As mentioned in 1.3.4.3, endoglucanase *cel5A* from *T. maritima* was successfully secreted in *G. thermoglucosidasius* [55], which indicates it can be a cellulase producer. By combination of biofuel producer, *G. thermoglucosidasius* is a potential cell factory for cellulosic biofuel from cellulose.

1.4 Consolidated bioprocessing

The processing strategy of cellulosic biofuel includes three steps (reviewed in 1.2.1): pretreatment, scarification and fermentation. These separate steps make the process inefficient, energy-intensive, and costly, which limits its application of being an alternative of gasoline. Therefore, great efforts have been made to simplify the process.

Biomass processing technology trends toward increasing consolidation over time. [153] In 1988, Wright *et al.* developed simultaneous saccharification and fermentation (SSF), consolidating hydrolysis and fermentation of cellulose hydrolysis products into one process [154]. In 1996, Lynd [155] proposed a highly compact design, which was later termed consolidated bioprocessing (CBP). It combines enzyme production, hydrolysis, and fermentation into a single step. CBP is paid much attention for economical cellulosic ethanol production [156,157]. As mentioned previously (1.3.6), cellulases from thermophilic bacteria are beneficial for cellulose hydrolysis, so thermophilic bacteria are good candidates for CBP together with their advantages of high temperature growth.

Two strategies are applied for developing CBP-enabling organism: one is engineering natural cellulolytic microorganisms to be biofuel producer; the other is engineering biofuel producers to be cellulolytic. Both of strategies are also employed for developing thermophilic CBP-enabling bacteria. The thermophilic cellulolytic bacterium *C. thermocellum* (introduced in 1.3.6) has been developed to increase ethanol tolerance and yield by using metabolic engineering tools[151].

Caldicellulosiruptor bescii is considered as the most thermophilic cellulolytic bacterium, with the ability to utilizing a wide range of substrates (C5 and C6 sugars, cellulose, hemicellulose, lignocellulosic plant biomass). Different genetic tools are developed to engineer this strain for improved hydrogen or ethanol production [158]. The thermophilic biofuel-producing bacterium *G. thermoglucosidasius* was engineered for degrading lignocellulose by expressing cellulase and xylanase [43,55].

1.5 Project overview

This thesis focuses on design and engineering of thermophilic *G. thermoglucosidasius* as a cell factory to produce biofuels and chemicals (Figure 1.4). To achieve this, the project contains three parts: 1) biofuel production, including metabolic engineering and evolutionary engineering for improved ethanol production, and metabolic engineering and conditions optimization for *R*-BDO production; 2) protein secretion, screening different signal peptides for heterologous protein secretion of α-amylase and cellulase; 3) consolidated bioprocessing, utilizing the evolved strain for ethanol production on starch or cellulose through expressing amylase or cellulases with efficient signal peptides.

In the third chapter of this thesis, an evolutionary engineering strategy was employed to increase ethanol production in *G. thermoglucosidasius*. Lactate and formate production in the strain of *G. thermoglucosidasius* 95A1was eliminated to divert carbon source for ethanol production. However, it was unable to grow under microaerobic conditions. To recover the growth, the strain was serial adapted in the presence of acetate, which resulted in evolved strains capable of efficiently producing ethanol during growth on glucose or cellobiose. Two loss-of-function mutations in evolved strains were identified through genome sequencing and verified through disruption of their genes. Same strategy was employed in another strain of *G.*

thermoglucosidasius C56-YS93, but it didn't improve ethanol production. Due to low efficiency of electroporation in *G. thermoglucosidasius* C56-YS93, a new transformation method of conjugation was established, which worked in both of the strains with a high efficiency. Besides, a new antibiotic selection system of spectinomycin was constructed and applied for gene deletion in the strains.

In the fourth chapter of this thesis, *G. thermoglucosidasius* 95A1 was engineered to produce high-value chemical, enantiopure (*2R, 3R*)-butanediol. For efficient protein expression, different promoters from *Geobacillus* and *B. subtilis* were screened firstly by using a thermostable green fluorescence protein (gfpmut3*) as a reporter. Then a new BDO biosynthetic pathway was introduced to *G. thermoglucosidasius* 95A1, consisting of heterologous acetolactate synthase (*alsS*) and acetolactate decarboxylase (*alsD*). In order to enhance *R*-BDO production, different *alsS* and *alsD* genes from thermophiles were cloned and tested. As a result, the strain expressing acetolactate synthase from *B. subtilis* and acetolactate decarboxylase from *Streptococcus thermophilus* under the lactate dehydrogenase promoter from *G. thermodenitrificans*, had the highest *R*-BDO production. The fermentation conditions were optimized to increase the production furtherly. Different alcohol dehydrogenase (*adh*) were deleted separately to divert more carbon source to *R*-BDO, but it only increased the yield but not production.

In the fifth chapter of this thesis, *G. thermoglucosidasius* 95A1 was engineered for heterologous protein secretion. The top 25 signal peptides, predicted by SignalP 4.1, from *G. thermoglucosidasius* C56-YS93 were screened for heterologous secretion of thermostable α-amylase from *G. stearothermophilus* and cellulase from *B. subtilis* and *C. bescii*. The efficient signal peptides for different secretory proteins were compared.

In the sixth chapter of this thesis, *G. thermoglucosidasius* 95A1 was engineered for consolidated bioprocessing. By combining the results of Chapter 3 and Chapter 5, strains for consolidated bioprocessing for ethanol production were constructed by transforming amylase or cellulase plasmids with relevant optimal signal peptides into evolved strains from the Chapter 3. As a result, the evolved strain with α-amylase secretion had efficient ethanol production directly from soluble starch; the evolved strains with different cellulase secretion produced little ethanol form cellulose, most probably because of low cellulase activity in the medium. For future development of this project, different strategies can be employed to increase hydrolysis efficiency, such as conditions optimization, introducing new enzymes and inactivation of serine proteases.

1.6 Figures

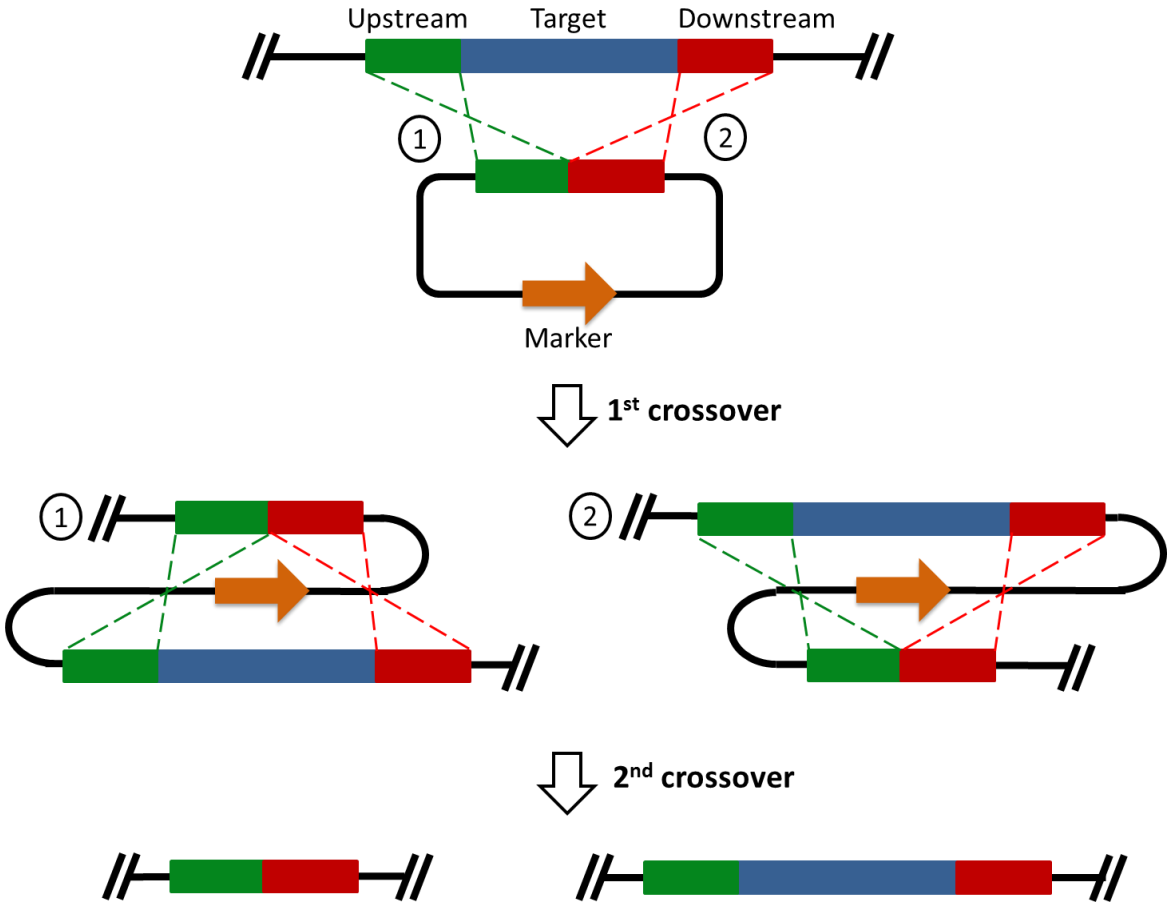

Figure 1.1 Gene knockout in *G. thermoglucosidasius*.

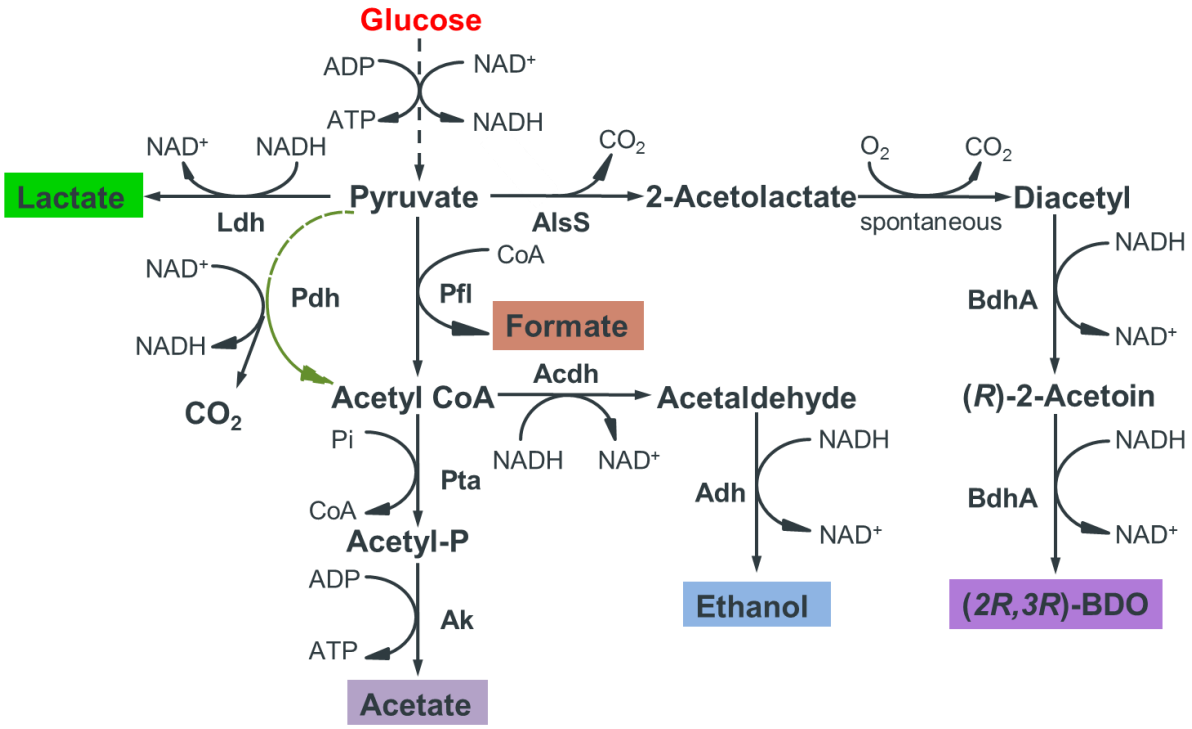

Figure 1.2 Schematic representation of fermentation pathways in *G. thermoglucosidasius*. Abbreviations: Ldh, lactate dehydrogenase; Pfl, pyruvate formate lyase; Pdh, pyruvate dehydrogenase; Acdh, acetaldehyde dehydrogenase; Pta, phosphotransacetylase; Ak, acetate kinase; Adh, alcohol dehydrogenase; AlsS, acetolactate synthase; BdhA, *2R*, *3R*-butanediol dehydrogenase; (*2R*, *3R*)-BDO, (*2R*, *3R*)-butanediol.

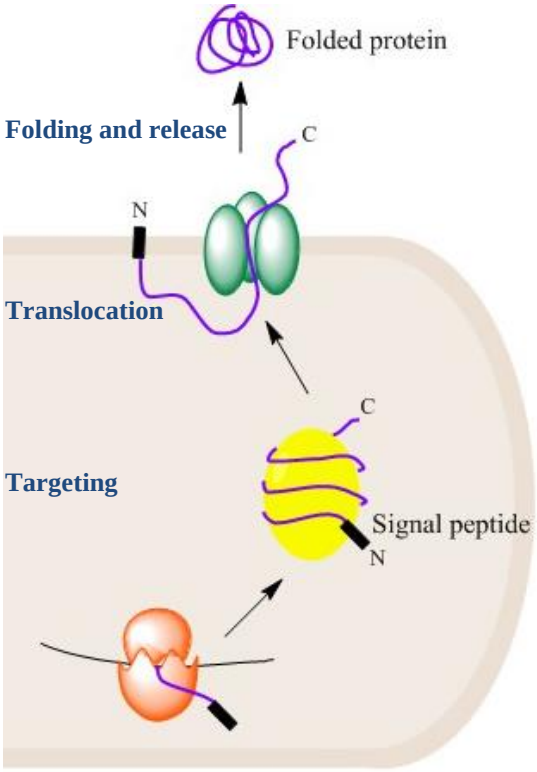

Figure 1.3 Sec pathway in *Geobacillus*.

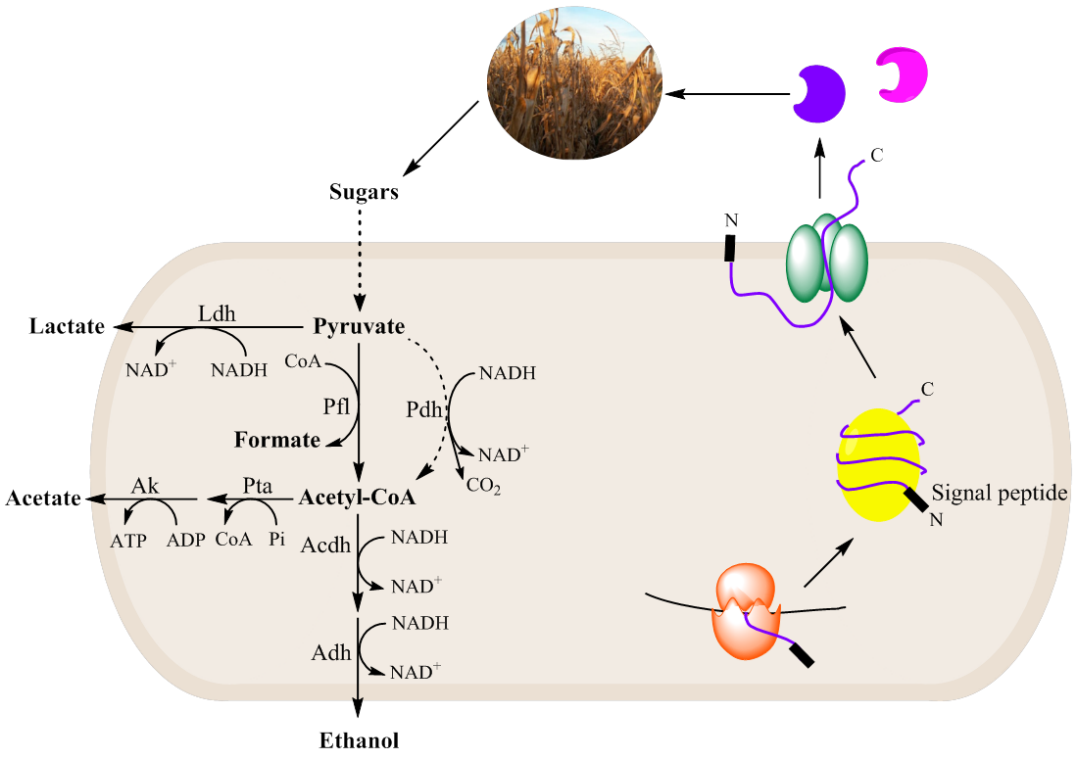

Figure 1.4 Produce fuels/chemicals from cellulosic biomass by *G. thermoglucosidasius*.

1.7 References

1. Kananavičiūtė R, Čitavičius D (2015) Genetic engineering of Geobacillus spp. Journal of Microbiological Methods 111: 31-39.
2. Pinzón‐Martínez D, Rodríguez‐Gómez C, Miñana‐Galbis D, Carrillo‐Chávez J, Valerio‐Alfaro G, et al. (2010) Thermophilic bacteria from Mexican thermal environments: isolation and potential applications. Environmental technology 31: 957-966.
3. Takaku H, Kodaira S, Kimoto A, Nashimoto M, Takagi M (2006) Microbial communities in the garbage composting with rice hull as an amendment revealed by culture-dependent and-independent approaches. Journal of bioscience and bioengineering 101: 42-50.
4. Strom PF (1985) Effect of temperature on bacterial species diversity in thermophilic solid-waste composting. Applied and Environmental Microbiology 50: 899-905.
5. Nazina TN, Tourova TP, Poltaraus AB, Novikova EV, Grigoryan AA, et al. (2001) Taxonomic study of aerobic thermophilic bacilli: descriptions of Geobacillus subterraneus gen. nov., sp. nov. and Geobacillus uzenensis sp. nov. from petroleum reservoirs and transfer of Bacillus stearothermophilus, Bacillus thermocatenulatus, Bacillus thermoleovorans, Bacillus kaustophilus, Bacillus thermodenitrificans to Geobacillus as the new combinations G. stearothermophilus, G. th. International Journal of Systematic and Evolutionary Microbiology 51: 433-446.
6. Zeigler DR (2005) Application of a recN sequence similarity analysis to the identification of species within the bacterial genus Geobacillus. International Journal of Systematic and Evolutionary Microbiology 55: 1171-1179.
7. Rahman T, Marchant R, Banat I (2004) Distribution and molecular investigation of highly thermophilic bacteria associated with cool soil environments. Portland Press Limited.
8. Takami H, Nishi S, Lu J, Shimamura S, Takaki Y (2004) Genomic characterization of thermophilic Geobacillus species isolated from the deepest sea mud of the Mariana Trench. Extremophiles 8: 351-356.
9. Hussein AH, Lisowska BK, Leak DJ (2015) The Genus Geobacillus and Their Biotechnological Potential. 92: 1-48.
10. Aliyu H, Lebre P, Blom J, Cowan D, De Maayer P (2016) Phylogenomic re-assessment of the thermophilic genus Geobacillus. Systematic and Applied Microbiology 39: 527-533.
11. Suzuki H, Murakami A, Yoshida Ki (2012) Counterselection System for Geobacillus kaustophilus HTA426 through Disruption of pyrF and pyrR. Applied and Environmental Microbiology 78: 7376-7383.
12. Cripps RE, Eley K, Leak DJ, Rudd B, Taylor M, et al. (2009) Metabolic engineering of Geobacillus thermoglucosidasius for high yield ethanol production. Metabolic Engineering 11: 398-408.
13. Lin PP, Rabe KS, Takasumi JL, Kadisch M, Arnold FH, et al. (2014) Isobutanol production at elevated temperatures in thermophilic Geobacillus thermoglucosidasius. Metabolic Engineering 24: 1-8.
14. Smerilli M, Neureiter M, Wurz S, Haas C, Frühauf S, et al. (2015) Direct fermentation of potato starch and potato residues to lactic acid by Geobacillus stearothermophilus under non‐sterile conditions. Journal of Chemical Technology and Biotechnology 90: 648-657.
15. Arena A, Gugliandolo C, Stassi G, Pavone B, Iannello D, et al. (2009) An exopolysaccharide produced by Geobacillus thermodenitrificans strain B3-72: antiviral activity on immunocompetent cells. Immunology letters 123: 132-137.
16. Xiao Z, Wang X, Huang Y, Huo F, Zhu X, et al. (2012) Thermophilic fermentation of acetoin and 2, 3-butanediol by a novel Geobacillus strain. Biotechnology for Biofuels 5: 88.
17. Ezeji TC, Bahl H (2006) Purification, characterization, and synergistic action of phytate-resistant α-amylase and α-glucosidase from Geobacillus thermodenitrificans HRO10. Journal of Biotechnology 125: 27-38.
18. Khemakhem B, Ali MB, Aghajari N, Juy M, Haser R, et al. (2009) Engineering of the α‐amylase from Geobacillus stearothermophilus US100 for detergent incorporation. Biotechnology and Bioengineering 102: 380-389.
19. Li H, Zhang X (2005) Characterization of thermostable lipase from thermophilic Geobacillus sp. TW1. Protein expression and purification 42: 153-159.
20. Soliman NA, Knoll M, Abdel-Fattah YR, Schmid RD, Lange S (2007) Molecular cloning and characterization of thermostable esterase and lipase from Geobacillus thermoleovorans YN isolated from desert soil in Egypt. Process Biochemistry 42: 1090-1100.

21. Zhu W, Cha D, Cheng G, Peng Q, Shen P (2007) Purification and characterization of a thermostable protease from a newly isolated Geobacillus sp. YMTC 1049. Enzyme and Microbial Technology 40: 1592-1597.
22. Sookkheo B, Sinchaikul S, Phutrakul S, Chen S-T (2000) Purification and characterization of the highly thermostable proteases from Bacillus stearothermophilus TLS33. Protein expression and purification 20: 142-151.
23. Ewis HE, Abdelal AT, Lu C-D (2004) Molecular cloning and characterization of two thermostable carboxyl esterases from Geobacillus stearothermophilus. Gene 329: 187-195.
24. Yildirim M, Colak A, Col M, Canakci S (2009) A new recombinant phosphotriesterase homology protein from Geobacillus caldoxylosilyticus TK4: An extremely thermo-and pH-stable esterase. Process Biochemistry 44: 1366-1373.
25. Sandalli C, Singh K, Modak M (2012) Characterization of catalytic carboxylate triad in non-replicative DNA polymerase III (pol E) of Geobacillus kaustophilus HTA. Cellular and molecular biology (Noisy-le-Grand, France) 58: 44-49.
26. Vellore J, Moretz SE, Lampson BC (2004) A group II intron-type open reading frame from the thermophile Bacillus (Geobacillus) stearothermophilus encodes a heat-stable reverse transcriptase. Applied and Environmental Microbiology 70: 7140-7147.
27. Tomita K, Nakajima T, Kikuchi Y, Miwa N (2004) Degradation of poly (L-lactic acid) by a newly isolated thermophile. Polymer Degradation and Stability 84: 433-438.
28. Banat IM, Marchant R (2011) Geobacillus activities in soil and oil contamination remediation. Endospore-forming Soil Bacteria: Springer. pp. 259-270.
29. Feng L, Wang W, Cheng J, Ren Y, Zhao G, et al. (2007) Genome and proteome of long-chain alkane degrading Geobacillus thermodenitrificans NG80-2 isolated from a deep-subsurface oil reservoir. Proceedings of the National Academy of Sciences 104: 5602-5607.
30. Shintani M, Ohtsubo Y, Fukuda K, Hosoyama A, Ohji S, et al. (2014) Complete genome sequence of the thermophilic polychlorinated biphenyl degrader Geobacillus sp. strain JF8 (NBRC 109937). Genome announcements 2: e01213-01213.
31. Mohammed Fayaz A, Girilal M, Rahman M, Venkatesan R, Kalaichelvan PT (2011) Biosynthesis of silver and gold nanoparticles using thermophilic bacterium Geobacillus stearothermophilus. Process Biochemistry 46: 1958-1962.
32. Taylor MP, van Zyl L, Tuffin IM, Leak DJ, Cowan DA (2011) Genetic tool development underpins recent advances in thermophilic whole-cell biocatalysts. Microbial Biotechnology 4: 438-448.
33. Imanaka T, Fujii M, Aramori I, Aiba S (1982) Transformation of Bacillus stearothermophilus with plasmid DNA and characterization of shuttle vector plasmids between Bacillus stearothermophilus and Bacillus subtilis. Journal of Bacteriology 149: 824-830.
34. Wu L, Welker N (1989) Protoplast transformation of Bacillus stearothermophilus NUB36 by plasmid DNA. Microbiology 135: 1315-1324.
35. Liao H, McKenzie T, Hageman R (1986) Isolation of a thermostable enzyme variant by cloning and selection in a thermophile. Proceedings of the National Academy of Sciences 83: 576-580.
36. Narumi I, Nakayama N, Nakamoto S, Kimura T, Yanagisawa T, et al. (1993) Construction of a new shuttle vector pSTE33 and its stabilities in Bacillus stearothermophilus, Bacillus subtilis, and Escherichia coli. Biotechnology Letters 15: 815-820.
37. Narumi I, Sawakami K, Nakamoto S, Nakayama N, Yanagisawa T, et al. (1992) A newly isolatedBacillusstearotheromophilus K1041 and its transformation by electroporation. Biotechnology techniques 6: 83-86.
38. Lu YP, Zhang C, Lv F, Bie X, Lu ZX (2012) Study on the electro‐transformation conditions of improving transformation efficiency for Bacillus subtilis. Letters in applied microbiology 55: 9-14.
39. Weaver JC, Chizmadzhev YA (1996) Theory of electroporation: a review. Bioelectrochemistry and bioenergetics 41: 135-160.
40. Zeigler DR (2001) The genus Geobacillus. Introduction and strain catalog. Catalog of strains 3.
41. Kananavičiūtė R, Butaitė E, Čitavičius D (2014) Characterization of two novel plasmids from Geobacillus sp. 610 and 1121 strains. Plasmid 71: 23-31.
42. Taylor MP, Esteban CD, Leak DJ (2008) Development of a versatile shuttle vector for gene expression in Geobacillus spp. Plasmid 60: 45-52.
43. Fong SS, Pogrebnyakov I, Jendresen CB, Nielsen AT (2017) Genetic toolbox for controlled expression of functional proteins in Geobacillus spp. PloS one 12: e0171313.
44. Holmes R, Jobling M (1996) Genetics: conjugation. Barons medical microbiology 4.

45. Suzuki H (2012) Genetic Transformation of Geobacillus kaustophilus HTA426 by Conjugative Transfer of Host-Mimicking Plasmids. Journal of Microbiology and Biotechnology 22: 1279-1287.
46. Johnsborg O, Eldholm V, Håvarstein LS (2007) Natural genetic transformation: prevalence, mechanisms and function. Research in microbiology 158: 767-778.
47. Blomqvist T, Steinmoen H, Håvarstein LS (2006) Natural genetic transformation: a novel tool for efficient genetic engineering of the dairy bacterium Streptococcus thermophilus. Applied and Environmental Microbiology 72: 6751-6756.
48. Koyama Y, Hoshino T, Tomizuka N, Furukawa K (1986) Genetic transformation of the extreme thermophile Thermus thermophilus and of other Thermus spp. Journal of Bacteriology 166: 338-340.
49. Shaw AJ, Hogsett DA, Lynd LR (2010) Natural Competence in Thermoanaerobacter and Thermoanaerobacterium Species. Applied and Environmental Microbiology 76: 4713-4719.
50. Mirończuk AM, Kovács ÁT, Kuipers OP (2008) Induction of natural competence in Bacillus cereus ATCC14579. Microbial Biotechnology 1: 226-235.
51. Gryczan T, Contente S, Dubnau D (1978) Characterization of Staphylococcus aureus plasmids introduced by transformation into Bacillus subtilis. Journal of Bacteriology 134: 318-329.
52. Nakayama N, Narumi I, Nakamoto S, Kihara H (1993) Complete nucleotide sequence of pSTK1, a cryptic plasmid from Bacillus stearothermophilus TK015. Biotechnology Letters 15: 1013-1016.
53. Stuknyte M, Guglielmetti S, Mora D, Kuisiene N, Parini C, et al. (2008) Complete nucleotide sequence of pGS18, a 62.8-kb plasmid from Geobacillus stearothermophilus strain 18. Extremophiles 12: 415-429.
54. Liao HH, Kanikula AM (1990) Increased efficiency of transformation ofBacillus stearothermophilus by a plasmid carrying a thermostable kanamycin resistance marker. Current microbiology 21: 301-306.
55. Bartosiak-Jentys J, Hussein AH, Lewis CJ, Leak DJ (2013) Modular system for assessment of glycosyl hydrolase secretion in Geobacillus thermoglucosidasius. Microbiology 159: 1267-1275.
56. Sheng L, Kovács K, Winzer K, Zhang Y, Minton NP (2017) Development and implementation of rapid metabolic engineering tools for chemical and fuel production in Geobacillus thermoglucosidasius NCIMB 11955. Biotechnology for Biofuels 10.
57. Bacon LF, Hamley-Bennett C, Danson MJ, Leak DJ (2017) Development of an efficient technique for gene deletion and allelic exchange in Geobacillus spp. Microbial Cell Factories 16.
58. Dale BE, Anderson JE, Brown RC, Csonka S, Dale VH, et al. (2014) Take a Closer Look: Biofuels Can Support Environmental, Economic and Social Goals. Environmental Science & Technology 48: 7200-7203.
59. Solomon BD (2010) Biofuels and sustainability. ANNALS OF THE NEW YORK ACADEMY OF SCIENCES 1185: 119-134.
60. Houghton J, Weatherwax S, Ferrell J (2006) Breaking the biological barriers to cellulosic ethanol: a joint research agenda. EERE Publication and Product Library.
61. Kim SR, Ha S-J, Wei N, Oh EJ, Jin Y-S (2012) Simultaneous co-fermentation of mixed sugars: a promising strategy for producing cellulosic ethanol. Trends in Biotechnology 30: 274-282.
62. Hong K-K, Nielsen J (2012) Metabolic engineering of Saccharomyces cerevisiae: a key cell factory platform for future biorefineries. Cellular and Molecular Life Sciences 69: 2671-2690.
63. Ha S-J, Galazka JM, Kim SR, Choi J-H, Yang X, et al. (2011) Engineered Saccharomyces cerevisiae capable of simultaneous cellobiose and xylose fermentation. Proceedings of the National Academy of Sciences 108: 504-509.
64. Kim SR, Park Y-C, Jin Y-S, Seo J-H (2013) Strain engineering of Saccharomyces cerevisiae for enhanced xylose metabolism. Biotechnology advances 31: 851-861.
65. Munjal N, Jawed K, Wajid S, Yazdani SS (2015) A Constitutive Expression System for Cellulase Secretion in Escherichia coli and Its Use in Bioethanol Production. PloS one 10: e0119917.
66. Dunn KL, Rao CV (2015) High-throughput sequencing reveals adaptation-induced mutations in pentose-fermenting strains ofZymomonas mobilis. Biotechnology and Bioengineering: n/a-n/a.
67. Xu Z, Huang F (2014) Pretreatment methods for bioethanol production. Applied biochemistry and biotechnology 174: 43-62.
68. Chang T, Yao S (2011) Thermophilic, lignocellulolytic bacteria for ethanol production: current state and perspectives. Applied Microbiology and Biotechnology 92: 13-27.
69. Olson DG, Sparling R, Lynd LR (2015) Ethanol production by engineered thermophiles. Current Opinion in Biotechnology 33: 130-141.
70. Olson DG, McBride JE, Shaw AJ, Lynd LR (2012) Recent progress in consolidated bioprocessing. Current Opinion in Biotechnology 23: 396-405.

71. Raita M, Ibenegbu C, Champreda V, Leak DJ (2016) Production of ethanol by thermophilic oligosaccharide utilising Geobacillus thermoglucosidasius TM242 using palm kernel cake as a renewable feedstock. Biomass and Bioenergy 95: 45-54.
72. Thompson AH, Studholme DJ, Green EM, Leak DJ (2008) Heterologous expression of pyruvate decarboxylase in Geobacillus thermoglucosidasius. Biotechnology Letters 30: 1359-1365.
73. Celińska E, Grajek W (2009) Biotechnological production of 2,3-butanediol—Current state and prospects. Biotechnology advances 27: 715-725.
74. Ji X-J, Huang H, Ouyang P-K (2011) Microbial 2, 3-butanediol production: a state-of-the-art review. Biotechnology advances 29: 351-364.
75. Ji X-J, Huang H, Ouyang P-K (2011) Microbial 2,3-butanediol production: A state-of-the-art review. Biotechnology advances 29: 351-364.
76. Ji XJ, Liu LG, Shen MQ, Nie ZK, Tong YJ, et al. (2015) Constructing a synthetic metabolic pathway in Escherichia coli to produce the enantiomerically pure (R, R) ‐ 2, 3 ‐ butanediol. Biotechnology and Bioengineering 112: 1056-1059.
77. Yang T, Rao Z, Hu G, Zhang X, Liu M, et al. (2015) Metabolic engineering of Bacillus subtilis for redistributing the carbon flux to 2,3-butanediol by manipulating NADH levels. Biotechnology for Biofuels 8.
78. Lian J, Chao R, Zhao H (2014) Metabolic engineering of a Saccharomyces cerevisiae strain capable of simultaneously utilizing glucose and galactose to produce enantiopure (2R,3R)-butanediol. Metabolic Engineering 23: 92-99.
79. Kim S-J, Seo S-O, Jin Y-S, Seo J-H (2013) Production of 2, 3-butanediol by engineered Saccharomyces cerevisiae. Bioresource technology 146: 274-281.
80. Xie N-Z, Chen X-R, Wang Q-Y, Chen D, Du Q-S, et al. (2017) Microbial Routes to (2R, 3R)-2, 3-Butanediol: Recent Advances and Future Prospects. Current Topics in Medicinal Chemistry 17: 2433-2439.
81. Yan Y, Lee C-C, Liao JC (2009) Enantioselective synthesis of pure (R, R)-2, 3-butanediol in Escherichia coli with stereospecific secondary alcohol dehydrogenases. Organic & biomolecular chemistry 7: 3914-3917.
82. Lian J, Chao R, Zhao H (2014) Metabolic engineering of a Saccharomyces cerevisiae strain capable of simultaneously utilizing glucose and galactose to produce enantiopure (2R, 3R)-butanediol. Metabolic Engineering 23: 92-99.
83. Kim S-J, Seo S-O, Jin Y-S, Seo J-H (2013) Production of 2,3-butanediol by engineered Saccharomyces cerevisiae. Bioresource technology 146: 274-281.
84. Gao J, Xu H, Li Q-j, Feng X-h, Li S (2010) Optimization of medium for one-step fermentation of inulin extract from Jerusalem artichoke tubers using Paenibacillus polymyxa ZJ-9 to produce R, R-2, 3-butanediol. Bioresource technology 101: 7076-7082.
85. Li L, Li K, Wang K, Chen C, Gao C, et al. (2014) Efficient production of 2, 3-butanediol from corn stover hydrolysate by using a thermophilic Bacillus licheniformis strain. Bioresource technology 170: 256-261.
86. Li L, Chen C, Li K, Wang Y, Gao C, et al. (2014) Efficient simultaneous saccharification and fermentation of inulin to 2, 3-butanediol by thermophilic Bacillus licheniformis ATCC 14580. Applied and Environmental Microbiology 80: 6458-6464.
87. Jia X, Peng X, Liu Y, Han Y (2017) Conversion of cellulose and hemicellulose of biomass simultaneously to acetoin by thermophilic simultaneous saccharification and fermentation. Biotechnology for Biofuels 10: 232.
88. Wang Q, Chen T, Zhao X, Chamu J (2012) Metabolic engineering of thermophilic Bacillus licheniformis for chiral pure D ‐ 2, 3 ‐ butanediol production. Biotechnology and Bioengineering 109: 1610-1621.
89. Gollop N, Damri B, Chipman D, Barak Z (1990) Physiological implications of the substrate specificities of acetohydroxy acid synthases from varied organisms. Journal of Bacteriology 172: 3444-3449.
90. Atsumi S, Hanai T, Liao JC (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. nature 451: 86-89.
91. Holland IB (2004) Translocation of bacterial proteins—an overview. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1694: 5-16.
92. Harwood CR, Cranenburgh R (2008) Bacillus protein secretion: an unfolding story. Trends in Microbiology 16: 73-79.
93. Tjalsma H, Antelmann H, Jongbloed JDH, Braun PG, Darmon E, et al. (2004) Proteomics of Protein Secretion by Bacillus subtilis: Separating the "Secrets" of the Secretome. Microbiology and molecular biology reviews 68: 207-233.
94. Ling Lin F, Zi Rong X, Wei Fen L, Jiang Bing S, Ping L, et al. (2007) Protein secretion pathways in Bacillus subtilis: Implication for optimization of heterologous protein secretion. Biotechnology advances 25: 1-12.

95. van Roosmalen ML, Geukens N, Jongbloed JD, Tjalsma H, Dubois J-YF, et al. (2004) Type I signal peptidases of Gram-positive bacteria. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1694: 279-297.
96. Tjalsma H, Bolhuis A, Jongbloed JD, Bron S, van Dijl JM (2000) Signal peptide-dependent protein transport inBacillus subtilis: a genome-based survey of the secretome. Microbiology and molecular biology reviews 64: 515-547.
97. van Dijl JM, Braun PG, Robinson C, Quax WJ, Antelmann H, et al. (2002) Functional genomic analysis of the Bacillus subtilis Tat pathway for protein secretion. Journal of Biotechnology 98: 243-254.
98. Sargent F, Bogsch EG, Stanley NR, Wexler M, Robinson C, et al. (1998) Overlapping functions of components of a bacterial Sec‐independent protein export pathway. The EMBO Journal 17: 3640-3650.
99. Hynds PJ, Robinson D, Robinson C (1998) The sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane. Journal of Biological Chemistry 273: 34868-34874.
100. Berks BC, Sargent F, Palmer T (2000) The Tat protein export pathway. Molecular microbiology 35: 260-274.
101. Palmer T, Berks BC (2012) The twin-arginine translocation (Tat) protein export pathway. Nature reviews Microbiology 10: 483.
102. Jongbloed JD, van der Ploeg R, van Dijl JM (2006) Bifunctional TatA subunits in minimal Tat protein translocases. Trends in Microbiology 14: 2-4.
103. Robinson C, Bolhuis A (2004) Tat-dependent protein targeting in prokaryotes and chloroplasts. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1694: 135-147.
104. von Heijne G (1990) The signal peptide. Journal of Membrane Biology 115: 195-201.
105. Kang Z, Yang S, Du G, Chen J (2014) Molecular engineering of secretory machinery components for high-level secretion of proteins in Bacillus species. Journal of industrial microbiology & biotechnology 41: 1599-1607.
106. Briggs MS, Cornell DG, Dluhy RA, Gierasch LM (1986) Conformations of signal peptides induced by lipids suggest initial steps in protein export. Science 233: 206-209.
107. Vrije Gd, Batenburg A, Killian J, Kruijff Bd (1990) Lipid involvement in protein translocation in Escherichia coli. Molecular microbiology 4: 143-150.
108. Pohl S, Harwood CR (2010) Heterologous protein secretion by Bacillus species: From the cradle to the grave. Advances in applied microbiology 73: 1-25.
109. Darmon E, Noone D, Masson A, Bron S, Kuipers OP, et al. (2002) A novel class of heat and secretion stress-responsive genes is controlled by the autoregulated CssRS two-component system of Bacillus subtilis. Journal of Bacteriology 184: 5661-5671.
110. Stephenson K, Harwood CR (1998) Influence of a cell-wall-associated protease on production of α-amylase by Bacillus subtilis. Applied and Environmental Microbiology 64: 2875-2881.
111. Westers H, Westers L, Darmon E, van Dijl JM, Quax WJ, et al. (2006) The CssRS two‐component regulatory system controls a general secretion stress response in Bacillus subtilis. The FEBS journal 273: 3816-3827.
112. Antelmann H, Darmon E, Noone D, Veening JW, Westers H, et al. (2003) The extracellular proteome of Bacillus subtilis under secretion stress conditions. Molecular microbiology 49: 143-156.
113. Krishnappa L, Dreisbach A, Otto A, Goosens VJ, Cranenburgh RM, et al. (2013) Extracytoplasmic Proteases Determining the Cleavage and Release of Secreted Proteins, Lipoproteins, and Membrane Proteins inBacillus subtilis. Journal of Proteome Research 12: 4101-4110.
114. Harwood CR (1992) Bacillus subtilis and its relatives: molecular biological and industrial workhorses. Trends in Biotechnology 10: 247-256.
115. Heng C, Chen Z, Du L, Lu F (2005) Expression and secretion of an acid-stable α-amylase gene in Bacillus subtilis by SacB promoter and signal peptide. Biotechnology Letters 27: 1731-1737.
116. Hatada Y, Hidaka Y, Nogi Y, Uchimura K, Katayama K, et al. (2004) Hyper-production of an isomalto-dextranase of an Arthrobacter sp. by a proteases-deficient Bacillus subtilis: sequencing, properties, and crystallization of the recombinant enzyme. Applied Microbiology and Biotechnology 65: 583-592.
117. Udaka S, Yamagata H (1993) [3] High-level secretion of heterologous proteins Bacillus brevis. Methods in enzymology 217: 23-33.
118. Yamagata H, Adachi T, Tsuboi A, Takao M, Sasaki T, et al. (1987) Cloning and characterization of the 5'region of the cell wall protein gene operon in Bacillus brevis 47. Journal of Bacteriology 169: 1239-1245.
119. Degering C, Eggert T, Puls M, Bongaerts J, Evers S, et al. (2010) Optimization of Protease Secretion in Bacillus subtilis and Bacillus licheniformis by Screening of Homologous and Heterologous Signal Peptides. Applied and Environmental Microbiology 76: 6370-6376.

120. Malten M, Nahrstedt H, Meinhardt F, Jahn D (2005) Coexpression of the type I signal peptidase gene sipM increases recombinant protein production and export in Bacillus megaterium MS941. Biotechnology and Bioengineering 91: 616-621.
121. Le Loir Y, Azevedo V, Oliveira SC, Freitas DA, Miyoshi A, et al. (2005) Protein secretion in Lactococcus lactis: an efficient way to increase the overall heterologous protein production. Microbial Cell Factories 4: 2.
122. Le Loir Y, Gruss A, Ehrlich S, Langella P (1998) A Nine-Residue Synthetic Propeptide Enhances Secretion Efficiency of Heterologous Proteins inLactococcus lactis. Journal of Bacteriology 180: 1895-1903.
123. Poquet I, Saint V, Seznec E, Simoes N, Bolotin A, et al. (2000) HtrA is the unique surface housekeeping protease in Lactococcus lactis and is required for natural protein processing. Molecular microbiology 35: 1042-1051.
124. Anné J, Vrancken K, Van Mellaert L, Van Impe J, Bernaerts K (2014) Protein secretion biotechnology in Gram-positive bacteria with special emphasis on Streptomyces lividans. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1843: 1750-1761.
125. Watanabe K, Tsuchida Y, Okibe N, Teramoto H, Suzuki N, et al. (2009) Scanning the Corynebacterium glutamicum R genome for high-efficiency secretion signal sequences. Microbiology 155: 741-750.
126. Bermúdez-Humarán LG, Kharrat P, Chatel J-M, Langella P (2011) Lactococci and lactobacilli as mucosal delivery vectors for therapeutic proteins and DNA vaccines. Microbial Cell Factories 10: S4.
127. Hemmerich J, Rohe P, Kleine B, Jurischka S, Wiechert W, et al. (2016) Use of a Sec signal peptide library from Bacillus subtilis for the optimization of cutinase secretion in Corynebacterium glutamicum. Microbial Cell Factories 15: 208.
128. Brockmeier U, Caspers M, Freudl R, Jockwer A, Noll T, et al. (2006) Systematic Screening of All Signal Peptides from Bacillus subtilis: A Powerful Strategy in Optimizing Heterologous Protein Secretion in Gram-positive Bacteria. Journal of Molecular Biology 362: 393-402.
129. Caspers M, Brockmeier U, Degering C, Eggert T, Freudl R (2010) Improvement of Sec-dependent secretion of a heterologous model protein in Bacillus subtilis by saturation mutagenesis of the N-domain of the AmyE signal peptide. Applied Microbiology and Biotechnology 86: 1877-1885.
130. Low KO, Mahadi NM, Illias RM (2013) Optimisation of signal peptide for recombinant protein secretion in bacterial hosts. Applied Microbiology and Biotechnology 97: 3811-3826.
131. Jonet MA, Mahadi NM, Murad AMA, Rabu A, Bakar FDA, et al. (2012) Optimization of a heterologous signal peptide by site-directed mutagenesis for improved secretion of recombinant proteins in Escherichia coli. Journal of molecular microbiology and biotechnology 22: 48-58.
132. Wu S-C, Yeung JC, Duan Y, Ye R, Szarka SJ, et al. (2002) Functional production and characterization of a fibrin-specific single-chain antibody fragment from Bacillus subtilis: effects of molecular chaperones and a wall-bound protease on antibody fragment production. Applied and Environmental Microbiology 68: 3261-3269.
133. Krishnappa L, Monteferrante CG, Neef J, Dreisbach A, van Dijl JM (2013) Degradation of Extracytoplasmic Catalysts for Protein Folding in Bacillus subtilis. Applied and Environmental Microbiology 80: 1463-1468.
134. Suzuki H, Yoshida K-i, Ohshima T (2013) Polysaccharide-degrading thermophiles generated by heterologous gene expression in Geobacillus kaustophilus HTA426. Applied and Environmental Microbiology 79: 5151-5158.
135. Holland A (2017) Optimisation of feedstock utilisation by Geobacillus thermoglucosidasius: University of Bath.
136. Noorwez S, Ezhilvannan M, Satyanarayana T (2006) Production of a high maltose-forming, hyperthermostable and Ca 2+-independent amylopullulanase by an extreme thermophile Geobacillus thermoleovorans in submerged fermentation.
137. Rao JUM, Satyanarayana T (2007) Improving production of hyperthermostable and high maltose-forming α-amylase by an extreme thermophile Geobacillus thermoleovorans using response surface methodology and its applications. Bioresource technology 98: 345-352.
138. Rao JUM, Satyanarayana T (2009) Hyperthermostable, Ca2+-Independent, and High Maltose-Forming α-Amylase Production by an Extreme Thermophile Geobacillus thermoleovorans: Whole Cell Immobilization. Applied biochemistry and biotechnology 159: 464-477.
139. Uma Maheswar Rao J, Satyanarayana T (2003) Enhanced secretion and low temperature stabilization of a hyperthermostable and Ca2+ - independent α - amylase of Geobacillus thermoleovorans by surfactants. Letters in applied microbiology 36: 191-196.

140. Wind R, Buitelaar R, Eggink G, Huizing H, Dijkhuizen L (1994) Characterization of a new Bacillus stearothermophilus isolate: a highly thermostable α-amylase-producing strain. Applied Microbiology and Biotechnology 41: 155-162.
141. Sivaramakrishnan S, Gangadharan D, Nampoothiri KM, Soccol CR, Pandey A (2006) a-Amylases from microbial sources–an overview on recent developments. Food Technol Biotechnol 44: 173-184.
142. Diderichsen B, Christiansen L (1988) Cloning of a maltogenic alpha-amylase from Bacillus stearothermophilus. FEMS microbiology letters 56: 53-59.
143. Cao H, Heel AJ, Ahmed H, Mols M, Kuipers OP (2017) Cell surface engineering of Bacillus subtilis improves production yields of heterologously expressed α-amylases. Microbial Cell Factories 16: 56.
144. Zarafeta D, Kissas D, Sayer C, Gudbergsdottir SR, Ladoukakis E, et al. (2016) Discovery and characterization of a thermostable and highly halotolerant GH5 cellulase from an icelandic hot spring isolate. PloS one 11: e0146454.
145. Liao H, Zhang X-Z, Rollin JA, Zhang Y-HP (2011) A minimal set of bacterial cellulases for consolidated bioprocessing of lignocellulose. Biotechnology Journal 6: 1409-1418.
146. Anbar M, Gul O, Lamed R, Sezerman UO, Bayer EA (2012) Improved thermostability of Clostridium thermocellum endoglucanase Cel8A by using consensus-guided mutagenesis. Applied and Environmental Microbiology 78: 3458-3464.
147. Acharya S, Chaudhary A (2012) Bioprospecting thermophiles for cellulase production: a review. Brazilian Journal of Microbiology 43: 844-856.
148. Johnson EA, Sakajoh M, Halliwell G, Madia A, Demain AL (1982) Saccharification of complex cellulosic substrates by the cellulase system from Clostridium thermocellum. Applied and Environmental Microbiology 43: 1125-1132.
149. Vazana Y, Moraïs S, Barak Y, Lamed R, Bayer EA (2010) Interplay between Clostridium thermocellum family 48 and family 9 cellulases in cellulosomal versus noncellulosomal states. Applied and Environmental Microbiology 76: 3236-3243.
150. Bien TLT, Tsuji S, Tanaka K, Takenaka S, Yoshida K-i (2014) Secretion of heterologous thermostable cellulases in Bacillus subtilis. The Journal of general and applied microbiology 60: 175-182.
151. Bhalla A, Bansal N, Kumar S, Bischoff KM, Sani RK (2013) Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. Bioresource technology 128: 751-759.
152. Viikari L, Alapuranen M, Puranen T, Vehmaanperä J, Siika-Aho M (2007) Thermostable enzymes in lignocellulose hydrolysis. Biofuels: 121-145.
153. Lynd LR, Weimer PJ, Van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiology and molecular biology reviews 66: 506-577.
154. Wright JD, Wyman CE, Grohmann K (1988) Simultaneous saccharification and fermentation of lignocellulose. Applied biochemistry and biotechnology 18: 75-90.
155. Lynd LR (1996) Overview and evaluation of fuel ethanol from cellulosic biomass: technology, economics, the environment, and policy. Annual review of energy and the environment 21: 403-465.
156. Zhang Y-HP, Lynd LR (2005) Cellulose utilization by Clostridium thermocellum: bioenergetics and hydrolysis product assimilation. Proceedings of the National Academy of Sciences of the United States of America 102: 7321-7325.
157. Van Zyl WH, Lynd LR, den Haan R, McBride JE (2007) Consolidated bioprocessing for bioethanol production using Saccharomyces cerevisiae. Biofuels: Springer. pp. 205-235.
158. Chung D, Cha M, Guss AM, Westpheling J (2014) Direct conversion of plant biomass to ethanol by engineered Caldicellulosiruptor bescii. Proceedings of the National Academy of Sciences 111: 8931-8936.

Chapter 2 Materials and Methods

2.1 Growth media

Luria-Bertani (LB, BD Difco^{TM}) media was used for growth of *E.coli* strains and SOC was used for recovery of cells for transformation.

TGP broth (17 g/L tryptone, 3 g/L soy peptone, 2.5 g/L K_2HPO_4 and 5 g/L NaCl supplemented with 4 g/L sodium pyruvate and 0.4% (v/v) glycerol after autoclaving) was used for routine liquid culture of *G. thermoglucosidasius*. Tryptose blood agar base (TBAB, BD Difco^{TM}) plates were used for transformation and routine solid culture. 2TY plates (16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, and 15 g/L agar) were used when selecting for plasmid integration.

Modified thermophilic minimal media TMLM (1 mg/L biotin, 1 mg/L thiamine · HCl, 1 mg/L nicotinic acid, 5 mg/L $CaCl_2$, 0.5 mg/L $FeCl_3 \cdot 6H_2O$, 5 mg/L $ZnSO_4 \cdot 7H_2O$, 10 mM $MnCl_2$, 1 g/L NH_4Cl, 1 g/L NaCl, 0.4 g/L $MgSO_4$, 12.5 g/L K_2HPO_4, 3 g/L KH_2PO_4, and 0.1% yeast extract) supplemented with glucose or cellobiose was used for ethanol production in test-tube fermentations. USYE media (10 mM $NaH_2PO_4 \cdot 2H_2O$, 10 mM K_2SO_4, 2mM citric acid, 1.25 mM $MgSO_4 \cdot 7H_2O$, 0.02 mM $CaCl_2 \cdot 2H_2O$, 1.65 mM $Na_2MoO_4 \cdot 2H_2O$, 20 mM urea, 25 µM $ZnSO_4 \cdot 7H_2O$, 100 µM $FeSO_4 \cdot 7H_2O$, 50 µM $MnSO_4 \cdot H_2O$, 5 µM $CuSO_4 \cdot 5H_2O$, 10 µM $CoSO_4 \cdot 7H_2O$, 16.85 µM $NiSO_4 \cdot 6H_2O$, 6.5 µM H_3BO_3 and 1% yeast extract supplemented with 12.5 µM biotin and glucose or cellobiose after autoclaving) was used for ethanol production in bioreactor fermentations.

Modified ASYE media (M9 medium, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1000X dilution of Trace Metal Mix A5 (2.86 g H_3BO_3, 1.81 g $MnCl_2 \cdot 4H_2O$, 0.222 g $ZnSO_4 \cdot 7H_2O$, 0.39 g $Na_2MoO_4 \cdot$

$2H_2O$, 0.079 g $CuSO_4 \cdot 5H_2O$, 49.4 mg $Co(NO_3)_2 \cdot 6H_2O$ per liter water), 0.01 g/l thiamin, 2 mM citric acid, 100 μM $FeSO_4 \cdot 7H_2O$, 16.85 μM $NiCl_3 \cdot 6H_2O$, 12.5 μM biotin and 0.2 M HEPES buffer) supplemented with glucose and yeast extract was used for *R*-BDO production in flask fermentations.

Antibiotics were used *Geobacillus* at the following concentrations: chloramphenicol, 7 μg/mL; kanamycin, 12.5 μg/mL and spectinomycin, 12.5 μg/mL (7 μg/mL for transformants selection).

2.2 Molecular biology

2.2.1 Gene cloning in *E.coli*

2.2.1.1 DNA amplification

To amplify the genes of interest, Phusion High-Fidelity DNA Polymerase or Q5 High-Fidelity DNA polymerase from New England Biolabs (NEB) were used for Polymerase chain reaction (PCR) reactions by following the recommended protocols on NEB website. To check cloning products, GoTaq Green Master Mix from Promega was used for colony PCR.

2.2.1.2 Cloning strategies

To construct plasmids, different cloning strategies were used including restriction enzyme cloning, TOPO cloning, Gibson assembly, Golden Gate assembly. Restriction enzymes and T4 DNA ligase for cloning were all from NEB and used according to their recommended protocols. The pGEM®-T Easy Vector Systems and Gibson assembly kit from Promega were used for TOPO cloning and two or three DNA fragments assembly, respectively, according to the manuals. The Golden Gate assembly was carried out in 15 μL of reaction as described [1],

including 25 ng of receiver plasmid, 75 ng of insert, 1 μL BsaI (NEB), and 1 μL T4 ligase and 1X ligase buffer. The reaction was performed in a thermocycler as follows: 37 ºC for 2 min, 16 ºC for 5 min, steps repeated 40-50 times, followed by 50 ºC for 5 min and 80 ºC for 5 min. The reaction product (5 μL) was transformed into DH5α competent cells.

2.2.1.3 Chemical transformation of *E.coli*

Preparation of competent cells by Inoue method

A 4 mL seed culture in LB was inoculated with a single colony from a fresh plate and grown aerobically at 37°C, 250 rpm for 6-8 hour. A culture in 50 mL LB was then inoculated (1:100 dilution) in a 250 mL flask and grown aerobically overnight (about 16h) at room temperature. When the OD_{600} of the culture reached 0.5, the cells were harvested by centrifugation for 10 min at 4°C and 4,000 rpm (Eppendorf centrifuge 5810R). The cells were then resuspended very gently in 20 mL cold Inoue buffer (55 mM $MnCl_2{\cdot}4H_2O$, 15 mM $CaCl_2{\cdot}2H_2O$, 250 mM KCl, 10 mM PIPES (use 0.5 M stock, pH 6.7)). The cells were centrifuged again and resuspended gently in 2 mL cold Inoue transformation buffer. The suspension was stored on ice for 10 min after 147 μL DMSO was added and mixed. It was dispensed in chilled, sterile microfuge tubes (250 μL each) and stored at -80 °C.

Transformation by heat-shock

5-10 μL DNA was mixed with 60 μL competent cells in a 1.5 mL chilled, sterile microfuge tube (5 μL for Gibson assembly or Golden Gate assembly products, 10 μL for ligation products). The tube was placed on ice for 30 min and then in water bath for 1.5 min at 42°C. It was put back on the ice for 3-5 min immediately after heat-shock. The cells were added with 1 mL SOC and recovered by shaking at 37°C for 1h. The cells were spread on plates with the appropriate

antibiotic and grown at 37°C. The recovery was carried out at 30°C for 2 h and the plates were put at 30°C, if colonies cannot grow well at 37°C (especially for protein expression plasmids construction).

2.2.2 Gene cloning in *Geobacillus*

2.2.2.1 Genomic DNA extraction

Genomic DNA of *Geobacillus* was prepared using the QIAamp DNA mini kit (Qiagen) following the recommended protocol for Gram-positive bacteria. For colony PCR, cells were resuspended in 20 μL TE buffer and incubated at 95°C for 10 min. The supernatant was then used as template for PCR.

2.2.2.2 Transformation

Electroporation

G. thermoglucosidasius 95A1 was transformed by electroporation using the method developed by Taylor and coworkers [2]. Frozen stocks were prepared firstly before making competent cells. Strain was grown overnight on a TBAB plate at 55°C or 60°C. A culture of 10 ml TGP (volumes can vary) was inoculated with a single colony and grown at 55°C, 250 rpm until OD_{600} reaches 0.8. The culture was added by an equal volume of 20% glycerol, divided into 1 ml aliquots and stored in 1.5ml tubes at -80°C. To compare competent cells, a culture of 50 ml TGP was then inoculated with 1 ml frozen stock in a 250 ml flask and incubate at 55° C, 250 rpm until OD_{600} reaches between 1.4 and 2.0, ideally 1.8. The culture was then cooled on ice for 10 minutes and the cells harvested by centrifuging at 4°C, 4,000 x *g*, 10 min. The cells were then washed three times using chilled electroporation buffer (0.5 M mannitol, 0.5 M sorbitol and 10%

glycerol). The volumes of the washing steps are 40 ml, 25 ml and 25 ml, respectively. The cells were then resuspended in 1 mL of chilled electroporation buffer and stored at -80°C. For each transformation, approximately 500 ng of plasmid DNA was mixed with 60 µl electro-competent cells. The mixture was then transferred to a chilled 1 mm electroporation cuvette and chilled on ice for 5 minutes. Electroporation was performed using a BioRad XCell gene pulser with the settings: 2.5 kV, 600 Ω, 10 µF. Following electroporation, 1 mL of pre-warmed (52°C) TGP broth was immediately added to the cuvette. The mixture was then transferred to a test tube and incubated at 52°C, 250 rpm for 1 hour. The cells were spread on TBAB plates with the appropriate antibiotic and incubated at 52°C for 24 h.

Conjugation

G. thermoglucosidasius C56-YS93 was transformed by conjugation based on a modified method [3]. Here, *E. coli* strain BW20767 was used as the donor and a Geobacillus-E.coli shuttle vector with *mob* was used as conjugative plasmid (the construction was described in 3.3.2). Firstly, BW20767 was transformed with conjugative plasmid carrying the desired gene by electroporation. Secondly, BW20767 containing the plasmid was grown in LB to mid-exponential phase (OD_{600} reached 0.5). Simultaneously, *G. thermoglucosidasius* C56-YS93 was also grown in TGP exponential phase (OD_{600} reached 0.8). Thirdly, the donor cells were then washed twice to remove the antibiotic, and resuspended gently in phosphate buffer saline. The two cell cultures were then mixed at a volumetric ratio of 1 donor to 5 recipients (the volume of BW20767 was 1 ml). Fourthly, the mixture was then filtered using a sterile nitrocellulose membrane (0.22 µm). The membrane was then placed on a TBAB plate without antibiotic and incubated overnight at 37°C. Fifthly, the cells were then resuspended in 1 mL of TGP media. A

100 μL volume (depend on efficiency) of cells were then spread on TBAB plates containing the appropriate antibiotic and incubated overnight at 52°C.

2.2.2.3 Gene knockout in Geobacillus

Genes were deleted in *G. thermoglucosidasius* 95A1 using the method previously developed by Cripps and coworkers [4]. The knockout plasmids were first transformed into *G. thermoglucosidasius* 95A1 by electroporation. The transformants were grown overnight in liquid 2TY media at 52°C and then grown 200-400 μL culture on 2TY plates at 68°C for about 24 h to induce single-cross plasmid integration. Stable, marker-free mutants were obtained by double, reciprocal crossover recombination. Single-cross integrants were then serially grown in TGP broth without antibiotics at 52°C for 16 h or 60°C for 8 h for 2-3 times. Single colonies were isolated from the last subculture by diluting (10^4-$4*10^4$) and spreading on TBAB plates without antibiotic. Colonies (usually 180 colonies were picked up) were then replica plated on TBAB plates with or without antibiotics. Colony PCR was used to distinguish between stable mutants and wild-type revertants. A similar method was used to delete genes in *G. thermoglucosidasius* C56-YS93, the difference being that conjugation was used instead of electroporation.

2.3 Analytical methods

2.3.1 HPLC

Sugar, lactate, formate, acetate and ethanol, *R*-BDO concentrations were measured using a Shimadzu high-performance liquid chromatography (HPLC) system equipped with an Aminex HPX-87H column (Bio-Rad) and Shimadzu RID-10A refractive index detector. Fermentation supernatants were first pretreated with trichloroacetic acid (TCA) to final concentration of 13% to precipitate protein (or with 3 volumes of methanol if the desired peak was covered by the one

of TCA). They were mixed well, incubated at 4°C for 30 min, and then centrifuged for 15 minutes at the highest speed. The resultant supernatant was then passed through 0.22 μm polyethersulfone syringe filter (or diluted 4 times, use the filter without dilution for methanol precipitation) before being analyzed by HPLC. The column was kept at 65°C, and 0.5 mM H_2SO_4 was used as a mobile phase at a constant flow rate of 0.5 mL/min. Samples (20 μl) were run for 30 minutes. Peaks were identified and quantified by retention time comparison to authentic standards.

2.3.2 Reduced sugar analysis

In order to evaluate enzyme activity of amylase and different cellulase, two broadly applicable reduced-sugar assays were used in a 96-well microplate format: DNS (Dinitrosalicylic acid) assay and pHBAH (*para*-hydroxybenzoic acid hydrazide) assay. The pHBAH assay is much more sensitive and less toxic than the DNS assay.

2.3.2.1 DNS method

The DNS solution was prepared as described [5]: 1% 3,5-Dinitrosalicylic acid, 1.6% sodium hydroxide, 30% potassium sodium tartrate. The reduced sugar can convert 3, 5-dinitrosalicylic acid (yellow) to 3-amino-5-nitrosalicylic acid (brown), which results in a change of light absorbance at wavelength of 540 nm. To estimate the concentration of reduced sugar, 40 μL of DNS solution (store at 4°C) was added to the same amount of samples in PCR tubes. The mixtures were then incubated at 95°C for 5 min on a thermocycler. After the cycler cooling down, 200 μL of iced water was added and mixed well. 100 μL of the mixtures were then transferred to a 96-well plate and measured the absorbance at 540nm [6]. D-glucose was used as a standard.

2.3.2.2 pHBAH assay

The pHBAH stock solution was prepared as described [7]: 5% w/v pHBAH, 0.5 M HCl, filtered through 0.8 μm filter, stored at 4°C. The working solution was prepared freshly before every measurement by mixing 1 volume of the stock solution with 4 volumes of 0.5 M NaOH. To estimate the concentration of reduced sugar, 120 μL of the working solution was added to 40 μL sample in PCR tubes. The mixtures were then incubated at 95°C for 5 min on a thermocycler. After the cycler cooling down, 100 μL of the mixture was then transferred to a 96-well plate and measured the absorbance at 410nm. D-glucose was used as a standard.

2.3.3 PASC preparation

Phosphoric acid swollen cellulose (PASC) is the model substrate for cellulase or cellulose consuming organism, which is generated from crystalline cellulose. The preparation was as follows [8]: 3 mL ddH_2O was slowly added to 1 g Avicel PH-101(Sigma) with stirring to form a homogeneous white paste. A volume of 25 mL iced 85% phosphoric acid was then added into the paste very slowly at 4°C, followed by the addition of another 25 mL phosphoric acid. The resulting solution was transparent and left at 4°C for 6 h or overnight with gentle stirring. To precipitate cellulose, 50 mL iced ddH_2O was added to the solution followed by vigorous stirring for 4 times (addition of 200 mL in total). The solution turned cloudy and was centrifuged at 4200 rpm, 4°C for 20 min. The pellet was washed with iced ddH_2O until pH was around 6. To remove residual phosphoric acid, the pellet was washed with 1% $NaHCO_3$ for three times and then with iced ddH_2O until pH was between 5 and 7.

2.3.4 CMC plates

The carboxymethyl cellulose (CMC) plates combined with Congo red staining were used to qualify secreted endoglucanase from different strains. The CMC plates were prepared by the addition of 2% CMC into TBAB or TMLM (with 1.5% agar). The culture of strains were spotted on the plates and incubated at 55°C overnight or until sufficient growth. The plate was firstly rinsed with ddH_2O to remove the plaques of strains, and stained with 0.1% Congo red (filtered through 0.22 μm filter) for 20 min. It was rinsed again and then de-stained with 1 M NaOH solution for 5-10 min by gently shaking. The de-staining was continued with new solution until the hydrolysis zones were clear.

2.4 References

1. Engler C, Gruetzner R, Kandzia R, Marillonnet S (2009) Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PloS one 4: e5553.
2. Taylor MP, Esteban CD, Leak DJ (2008) Development of a versatile shuttle vector for gene expression in Geobacillus spp. Plasmid 60: 45-52.
3. Suzuki H (2012) Genetic Transformation of Geobacillus kaustophilus HTA426 by Conjugative Transfer of Host-Mimicking Plasmids. Journal of Microbiology and Biotechnology 22: 1279-1287.
4. Cripps RE, Eley K, Leak DJ, Rudd B, Taylor M, et al. (2009) Metabolic engineering of Geobacillus thermoglucosidasius for high yield ethanol production. Metabolic Engineering 11: 398-408.
5. Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. Analytical chemistry 31: 426-428.
6. Cornelis P, Digneffe C, Willemot K (1982) Cloning and expression of a Bacillus coagulans amylase gene in Escherichia coli. Molecular and General Genetics MGG 186: 507-511.
7. Mellitzer A, Glieder A, Weis R, Reisinger C, Flicker K (2012) Sensitive high-throughput screening for the detection of reducing sugars. Biotechnology Journal 7: 155-162.
8. Zhang Y-HP, Cui J, Lynd LR, Kuang LR (2006) A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. Biomacromolecules 7: 644-648.

Chapter 3 Engineering *G. thermoglucosidasius* for Ethanol Production[1]

3.1 Introduction

Thermophilic *G. thermoglucosidasius* is a promising cellulosic biofuels and chemicals producer since the fermentation processes are less susceptible to contamination and potentially more energy efficient than those of mesophilic organism. In this project, we started from improving ethanol production in this strain, because it is well studied and widely used in industry. Multiple reports have recently demonstrated that diverse thermophilic bacteria can be engineered to produce ethanol with high yields by modulating the expression of native genes [1,2,3]. *G. thermoglucosidasius* natively produces a mixture of lactate, acetate, formate, and ethanol during fermentative growth on glucose or cellobiose [4]. To improve ethanol production in *G. thermoglucosidasius*, Cripps *et al*. previously deleted the genes encoding lactate dehydrogenase and pyruvate formate lyase and expressed the gene encoding pyruvate dehydrogenase by using the oxygen-insensitive promoter for lactate dehydrogenase from *G. stearothermophilus*. This design eliminated mixed-acid fermentation and resulted in ethanol being produced at high yields. In addition, a number of studies have explored design involving the expression of pyruvate decarboxylase from different heterologous hosts [5,6,7]. However, no increases in ethanol productivity were observed during growth at high (~60˚C) temperatures.

In this chapter, we employed an evolutionary metabolic engineering strategy to improve ethanol production in *G. thermoglucosidasius* 95A1. Our initial goal was to test whether

[1]Portions of this chapter were reused with permission from "Evolutionary engineering of *Geobacillus thermoglucosidasius* for improved ethanol production" in the journal Biotechnology and Bioengineering, 2016, by Zhou J, Wu K, Rao CV. This article is protected by copyright. All rights reserved.

expression of the gene encoding pyruvate decarboxylase from *Sarcina ventriculi* would improve ethanol production in a strain of *G. thermoglucosidasius* deleted for the gene encoding lactate dehydrogenase. This strain, however, produced significant formate. To eliminate formate production, we also deleted the gene encoding pyruvate formate lyase. The resulting strain grew poorly under microaerobic conditions, presumably because the cells were unable to produce acetyl-CoA. We found that the addition of acetate could partially rescue growth. We then serially adapted these strains on glucose and cellobiose in the presence of acetate. The evolved strains were able to grow fast on glucose or cellobiose and produce ethanol at high yields. By sequencing the evolved strains, we identified and verified the genes responsible for the improvement of growth and ethanol production. This chapter is adapted from my research article published in Biotechnology and Bioengineering [8].

3.2 Results

3.2.1 Construction of ethanol producing strain

Wild-type of *G. thermoglucosidasius* performs mixed-acid fermentation with little ethanol being produced during microaerobic growth on sugars (Figure 3.1 and 3.2). To improve ethanol production, the initial design involved deleting lactate dehydrogenase and then over-expressing pyruvate decarboxylase (*pdc*) from *Sarcina ventriculi* ATCC 55887 [9]. This design was motivated by the moderate thermostability (Tm~50-60°C) of the *S. ventriculi* pyruvate decarboxylase (data not shown).

As expected, we observed a 200% increase in final ethanol titers during growth on glucose in a strain where *ldh* was deleted (Figure 3.1a). However, formate titers also increased by nearly 100% due to the increased flux through pyruvate formate lyase and to the loss of the competing

pathway through lactate dehydrogenase (Figure 3.1f). The reason for the observation that the increases in ethanol are greater than the increases in formate are is likely due to the weak native expression of pyruvate dehyrdogenase under microaerobic conditions [4,10]. We next expressed the *pdc* from *S. ventriculi* using the native *recA* promoter from *G. thermoglucosidasius* 95A1 in the Δ*ldh* strain. No significant differences in final ethanol concentrations were observed in the P*recA::pdc* Δ*ldh* strain relative to the Δ*ldh* strain (Figure 3.1a), which is consistent with similar efforts by others [7,11].

We next investigated the effect of deleting *pfl* in *G. thermoglucosidasius* 95A1. However, the resulting strain (Δ*ldh* Δ*pfl*) grew very poorly under microaerobic conditions (Figure 3.3). Likely, this strain was unable to produce sufficient acetyl-CoA for growth, because acetyl-CoA is principally generated by pyruvate formate lyases in the absence of oxygen. To rescue growth in our Δ*ldh* Δpfl strain, we instead grew the cells in the presence of 0.1% acetic acid. The rational was that acetic acid would provide the cell with an alternate source for producing acetyl-CoA [12,13]. As predicted, the addition of acetic acid improved the growth and ethanol production (Figure 3.3a). The acetate concentration decreased during growth, which supported the assumption that addition of acetic acid drove the reversible conversion from acetate to acetyl-CoA (Figure 3.3d). However, the decrease was minor. We also tested higher concentrations of acetate, but found that they did not increase growth (Figure 3.3b).

3.2.2 Serial adaptation improves growth and ethanol production

Serial adaptation improves growth and ethanol production. The addition of acetic acid partially rescued growth of the Δ*ldh* Δ*pfl* strain but it was still poor. To improve the performance of this strain, we serially adapted the strain to growth in minimal media medium containing 0.1% acetic acid and 1% glucose or cellobiose using three parallel cultures. Glucose (cellobiose)

consumption and ethanol production significantly improved after the fifth transfer but changed little from the fifth transfer to the tenth (Figure 3.4a and 3.5a). Compared to the unadapted strain, the evolved strains exhibited significantly improved growth (Figure 3.4c and 3.5c) and ethanol production (Figure 3.4b and 3.5b). Evolution in cellobiose had similar increase on growth and ethanol production (data not shown). The ethanol yield increased from 0.38 to 0.43 (g ethanol/g glucose) (Table 3.1) and from 0.41 to 0.46 (g ethanol/g cellobiose) (Table 3.2), respectively. These results demonstrate that the evolved Δ*ldh* Δ*pfl* strain can produce ethanol at high yield. They are also comparable to the results obtained by Cripps and coworkers (0.42 for glucose and 0.47 for cellobiose).

The evolved strain still required acetic acid (Figure 3.6). As acetic acid was directly added to the growth medium without pH adjustment, resulting in a pH decrease from 7.3 to 6.9, we also test whether the pH drop alone explained the requirement (Figure 3.6). However, the pH adjustment alone had no effect, indicating that the presence of acetate was required.

We also tested whether expression of the *pdc* from *S. ventriculi* would improve ethanol production in the evolved strain (Figure 3.7). However, we observed no difference in final ethanol titers. These results further confirm our previous results regarding the *pdc*.

We also tested the performance of the evolved strains in bioreactors using 0.1% acetate and 3% glucose or cellobiose (Figure 3.8). Again, we observed robust growth though the ethanol yields (0.36 g ethanol/g glucose and 0.32 g ethanol/g cellobiose) were less than observed in test tubes. One possibility is that some ethanol evaporated and left in the off gas, something we did not measure unlike Cripps and coworkers [4]. Higher concentrations of sugars were also tried as well as operating the bioreactor in fed-batch mode. However, the final ethanol concentration never

exceeded 12.5 g/L. Likely, these higher ethanol concentrations are toxic to the cell. Indeed, we found complete growth arrest at ethanol concentrations greater than 10 g/L (Figure 3.9).

We also found that the cells produced acetate in the bioreactors during the later phases of growth on glucose (Figure 3.9). We found that air (0.05 vvm) needed to be added to the bioreactor in order to obtain robust growth. Acetate production is likely a consequence of aerobic metabolism, particularly when cell growth ceases.

3.2.3 Identification of mutations in the evolved strains

To determine the mutations that rescued growth in *G. thermoglucosidasius* 95A1, we sequenced the three parallel evolved strains (three on glucose and three on cellobiose). Fewer than 10 single nucleotide polymorphisms (SNPs) (Table 3.3) unique to each evolved strain relative to the parental strain were identified, and most of the SNPs were intergenic (non-coding regions) (Table 3.4). Among 7 genes containing non-synonymous SNPs, only mutations in the gene of *aprt* encoding adenine phosphoribosyltransferase was shared among all six evolved strains (Table 3.3). Adenine phosphoribosyltransferase is involved in the purine nucleotide salvage pathway, which catalyzes the reaction transforming adenine and phosphoribosyl pyrophosphate to adenosine monophosphate (AMP) and diphosphate (PPi) [14]. In the strains evolved in glucose, a single mutation was also found in the gene of *spoIIIAA* encoding stage III sporulation protein AA. The stage III sporulation protein AA is one of the SpoIIIA proteins, which forms a channel with SpoIIQ between the mother cell and forespore [15]. In the case of the *aprt* gene, two strains contained loss of function mutations (frame shift after Cys64) and the other four contained mutations of Phe74Val or Thr123Ile. In the case of the *spoIIIAA* gene, all three strains had the same mutation (frame shift after Ile4).

3.2.4 Effect of *aprt* and *spoIIIAA* on growth and ethanol production

We next tested whether deleting the *aprt* and *spoIIIAA* genes, separately and in combination, would improve growth and ethanol production in the unevolved Δ*ldh* Δ*pfl* strain. Deleting the *aprt* gene significantly improved growth on both glucose (Figure 3.10b) and cellobiose (Figure 3.11b) relative to the respective unevolved strains. While the increase was significant, it did not phenocopy the evolved strain. Deleting the *spoIIIAA* gene led to minor increase in ethanol production during growth on glucose (Figure 3.10a) but has no effect during growth on cellobiose (3.11a). Deleting both genes (Δ*ldh* Δ*pfl* Δ*aprt* Δ*spoIIIAA*) did not further improve the growth or ethanol production relative to the Δ*ldh* Δ*pfl* Δ*aprt* (Figure 3.10a). These results indicate that loss of *aprt* is a major factor governing the improved performance of the evolved strains whereas loss of the *spoIIIAA* gene had only a marginal effect.

3.2.5 Evolutionary metabolic engineering of *G. thermoglucosidasius* C56-YS93

One challenge in working with non-model organisms is that there is scant data to guide the choice of strain. Our choice to work with *G. thermoglucosidasius* 95A1 was based on the precedent established by others. However, no other strain had been explored to the best of our knowledge in the context of ethanol production. Therefore, we tested whether a similar approach could be employed using *G. thermoglucosidasius* C56-YS93. Unlike *G. thermoglucosidasius* 95A1, however, we were unable to transform *G. thermoglucosidasius* C56-YS93 using electroporation. We instead employed conjugation. Otherwise, the strategy was the same.

We were unable to improve ethanol production using identical growth conditions in G. thermoglucosidasius C56-YS93 (Table 3.5). Deletion of *ldh* in *G. thermoglucosidasius* C56-YS93 inhibited cell growth unlike the case with *G. thermoglucosidasius* 95A1. Further deletion of *pfl* and subsequent adaptation in the presence of acetate did not rescue growth. These results

demonstrate that *G. thermoglucosidasius* C56-YS93 is not amendable for metabolic engineering, at least in the case of ethanol production using the design employed.

3.3 Discussion

High-temperature fermentations are appealing for the production of fuels and other chemicals from lignocellulosic biomass, because they are potentially less susceptible to contamination and more energy efficient. Not surprisingly, many researchers have engineered thermophiles for the production of ethanol and other fuels [3,16,17,18,19,20]. In this study, we engineered *G. thermoglucosidasius* to efficiently produce ethanol from glucose and cellobiose using a combination of metabolic engineering and evolutionary engineering. This work builds on the pioneering work by Cripps and coworkers, who improved ethanol production in *G. thermoglucosidasius* by deleting the genes encoding lactate dehydrogenase and pyruvate formate lyase and over-expressing the native pyruvate dehydrogenase [4]. Our initial goal was to introduce an exogenous pathway using pyruvate decarboxylase from *S. ventriculi*, a Gram-positive bacterium, in strain deleted for lactate dehydrogenase. Heterologous expression of pyruvate decarboxylase from *Z. mobilis* [11] and *Gluconobacter oxydans* [7], both Gram-negative bacteria, was previously tested in *G. thermoglucosidasius*. Both were functional in. *G. thermoglucosidasius* at lower, sub-optimal growth temperatures around 50°C but were not functional at higher, more optimal growth temperatures around 60°C. Our initial hypothesis was that the pyruvate decarboxylase from *S. ventriculi* would be a better candidate as it exhibited moderate thermostability (data not shown) and came from a more closely related species of bacteria. However, we found that heterologous expression of pyruvate decarboxylase from *S. ventriculi* did not increase final ethanol titers during growth at 60°C, consistent with previous studies. We, therefore, explored an alternate strategy where we deleted the genes encoding

lactate dehydrogenase and pyruvate formate lyase. The cells, however, grew poorly during anaerobic growth and also did not produce ethanol.

The major finding in this work was that the addition of acetate partially rescues growth and improves ethanol production in the Δ*ldh* Δ*pfl* strain. While the exact mechanism is unknown, one possibility is that the strain is unable to produce sufficient quantities of acetyl-CoA from pyruvate in the absence of oxygen. Acetate provides an alternate source for making acetyl-CoA, which is necessary for both biosynthesis and regenerating NAD^+ (through the production of ethanol). In bacteria, there are two routes for producing acetyl-CoA from acetate: the high-affinity pathway involving acetyl-CoA synthetase (Acs) and the low-affinity one involving acetate kinase and phosphate acetyltransferase (Ack/Pta) [21]. Both are present in *G. thermoglucosidasius.*

From a fermentation standpoint, the acetate requirement is not problematic because it is present in lignocellulosic hydrolysates. Nominally, acetate is considered a nuisance, because it inhibits growth in many strains [22]. Indeed, significant work has focused on engineering acetate-tolerant microorganisms [3,23]. In our case, the presence of acetate is necessary for cell growth (Figure 3.6) and potentially represents an alternative strategy to engineering acetate tolerance.

Despite the ability of acetate to improve growth and ethanol production, both were still poor. We therefore employed an evolutionary approach to improve growth and ethanol production by serial adaptation in minimal media containing acetic acid and either glucose or cellobiose. After just a few transfers, we were able to isolate strains capable of efficiently producing ethanol from the two sugars. Genome sequencing of the evolved strains revealed common mutations in *aprt*, predicted to encode adenine phosphoribosyltransferase, during growth on either sugar. Common

mutations were also found in *spoIIIAA*, predicted to encode the stage III sporulation protein AA, during growth on glucose. Interestingly, no mutations were found in the gene or promoter of pyruvate dehydrogenase, which was contrary to what we initially anticipated based on the work of Cripps and coworkers.

Adenine phosphoribosyltransferase is involved in the purine nucleotide salvage pathway, which catalyzes the reaction transforming adenine and phosphoribosyl pyrophosphate to adenosine monophosphate (AMP) and diphosphate (PPi) [14]. Consistent with loss-of-function mutations in *aprt* being adaptive, deleting the gene in the unevolved Δ*ldh* Δ*pfl* strain increased the sugar consumption and ethanol production (Figure 3.10d and 3.11d). Acetate consumption increased also in this strain (Δ*ldh* Δ*pfl* Δ*aprt*), suggesting somehow that the loss of *aprt* increased the rate of conversion of acetate to acetyl-CoA, presumably via the Acs pathway. AMP and PPi are both products of the forward reactions catalyzed by Acs and adenine phosphoribosyltransferase. One possibility is that loss of adenine phosphoribosyltransferase could reduce amount of AMP and PPi in the cell, thus pushing acetate to acetyl-CoA via Acs by reducing the driving force for the reverse reaction. Such a mechanism would explain how an enzyme not involved in primary metabolism increases sugar consumption and ethanol production in *G. thermoglucosidasius*.

One caveat with the putative mechanism proposed above is that a second pathway exists for generating acetyl-CoA from acetate, the Ack/Pta pathway. While this pathway is nominally involved in acetate production from acetyl-CoA, it can also operate in the reverse direction when acetate concentrations are high [21], such as in our experiments. Unlike the Acs pathway, AMP and PPi are not involved in the Ack/Pta pathway. This raises the possibility that the growth defect is not due to acetyl-CoA per se but some other deficiency. One possibility is that the cells

are deficient in C1 metabolism due to their inability to produce formate, a precursor for formyl-tetrahydrofolate. Indeed, others have shown that Δ*pfl* mutants of *Staphylococcus aureus* and *C. thermocellum* grow poorly in the absence of oxygen, and that the addition of exogenous formate partially restores growth [24,25]. How loss-of-function mutations in *aprt* could affect C1 metabolism is not known. Nonetheless, we tested whether the addition of formate increased glucose consumption and ethanol production in the unevolved Δ*ldh* Δ*pfl* mutant (Figure 3.12). However, no increases were observed, indicating that C1 metabolism is not affected. Clearly, further work is necessary to determine exactly how acetate and *aprt* are affecting ethanol production in the Δ*ldh* Δ*pfl* mutant.

In Bacillus subtilis, the stage III sporulation protein AA is one of the SpoIIIA proteins, which forms a channel with SpoIIQ between the mother cell and forespore. The protein is similar to ATPases found in type II and IV secretion systems. The ATPase motif is essential for σG activation and efficient sporulation [15]. Deletion of *spoIIIAA* increased glucose consumption and ethanol production (Figure 3.10 and 3.11), indicating that a loss-of-function may be adaptive. However, the effect was minor. Furthermore, the same mutation was found in all three glucose-adapted strains, suggesting that it was acquired just prior to the glucose adaptation experiments (thought it was not present in parental strain or the cellobiose-adapted strains). This would also imply that the mutation is in fact not adaptive, and that the minor enhancement observed in the Δ*ldh* Δ*pfl* Δ*spoIIIAA* mutant resulted from secondary mutations acquired during strain construction or possibly the growth experiment itself. In addition, deletion of both *spoIIIAA* and *aprt* strain did not further increase sugar consumption and ethanol production beyond what was observed when *aprt* was deleted alone (Δ*ldh* Δ*pfl* Δ*aprt*), again suggesting that the *spoIIIAA*

mutant is neutral. As there is no evident connection, even tenuous ones, between SpoIIIAA and metabolism, we conclude that the *spoIIIAA* result is likely not significant.

Acetate concentrations did not decrease in the evolved strains (Figure 3.10 and 3.11). This was contrary to our expectation and our results concerning *aprt*. Somehow the adapted cells are able to make acetyl-CoA in the absence of oxygen and independent of acetate, even though acetic acid is still required for growth. Likely, these strains are making it from pyruvate. How they do this is not clear. Pyruvate formate lyase is not present and, as we noted already, no mutations related to pyruvate dehydrogenase were identified. Yet, activation of pyruvate dehydrogenase by some unknown mechanism provides the best explanation. Moreover, we found that expression of pyruvate decarboxylase from *S. ventriculi* had no effect in the evolved strain (Figure 3.7), again suggesting that the pyruvate dehydrogenase is active in the evolved strain.

Why then is acetic acid still required if the evolved cells can make acetyl-CoA directly from pyruvate? One possibility is that during the initial phases of growth, there is insufficient NAD^+ for glucose oxidation. By assimilating acetate, the cells can produce acetyl-CoA, which can regenerate NAD^+ through the production of ethanol. Indeed, we observe a small decrease in acetate concentrations during the early phases of growth for the evolved strain (Figure 3.10), followed by a commensurate increase such that final acetate concentrations are relatively unchanged. This would suggest that acetate is necessary to initiate robust growth but not to sustain it.

Our adaptation experiments were performed at relatively low sugar concentrations. Growth at higher sugar concentrations led to incomplete fermentations within sealed test tubes, even after adaptation. The reason is that pH decreases to approximate 6.0 during the fermentations, due to

the production of CO_2, which inhibits the growth of *G. thermoglucosidasius* (data not shown). We were able to perform fermentations at higher sugar concentrations using bioreactors where the pH was controlled. However, ethanol concentrations never exceeded 12.5 g/L due to toxicity. These results indicate that *G. thermoglucosidasius* 95A1 is not natively tolerant to ethanol and further engineering is required before this thermophile is capable of producing ethanol at high titers.

In addition to our work with *G. thermoglucosidasius* 95A1, we also explored whether a similar engineering strategy could be applied to G. *thermoglucosidasius* C56-YS93. Our results suggest that this strain is not a promising candidate for ethanol production, at least using the design employed here. In the course of our work with *G. thermoglucosidasius* C56-YS93, we found that it could not be transformed by electroporation unlike *G. thermoglucosidasius* 95A1 [4,11]. The only successful transformation method in our hands was conjugation, which is commonly used in *G. kaustophilus* [26]. We also found that conjugation can also be applied in *G. thermoglucosidasius* 95A1 and 95A2 (data not shown), indicating that it can be used for genetic manipulation in multiple species of *G. thermoglucosidasius.*

One challenge working with thermophiles is that the genetic tools are still limited. As a specific example, many antibiotics and their corresponding resistance proteins do not work due to limited thermostability. The most commonly used antibiotics in *Geobacilli* are kanamycin and chloramphenicol [27]. Kanamycin is more popular because it and its resistance protein, kanamycin nucleotidyltransferase, are active above 60°C [28]. In this work, we used spectinomycin for the first time in *Geobacillus* and found it was still selective at high temperature. The resistance marker (spc^R) was obtained by cloning *spc* gene for spectinomycin adenyltransferase from an *E. coli-Bacillus* shuttle vector pDG1664 [29]. It was used successfully

for both selection of transformants at 52°C and integrants at 68°C, which suggests it can be used not only in mesophilic bacteria but also in thermophilic bacteria.

3.4 Conclusion

In conclusion, we engineered *G. thermoglucosidasius* 95A1 to efficiently produce ethanol from glucose or cellobiose using a combination of metabolic and evolutionary engineering. This work further establishes this thermophile as a platform organism for fuel and chemical production. However, significant work is still required before *G. thermoglucosidasius* 95A1 is capable of producing ethanol and likely other fuels and chemicals at levels comparable to mesophilic microorganisms.

3.5 Material and methods

3.5.1 Strains and growth conditions

All strains used in this study are listed in Table 3.6. TGP broth was used for routine growth of *G. thermoglucosidasius*. Tryptose blood agar base (TBAB) plates were used for transformation. 2TY plates were used when selecting for plasmid integration. Modified thermophilic minimal medium supplemented with glucose or cellobiose was used for test-tube fermentations. USYE medium was used for bioreactor fermentations. Antibiotics were used at the following concentrations: chloramphenicol, 7 μg/mL; and spectinomycin, 7 μg/mL (12.5 μg/mL for selection of chromosomal integrants).

3.5.2 Plasmid construction

All plasmids used in this study are listed in Table 3.6. All primers used in this study are listed in Table 3.7. The pKW1012 shuttle plasmid was constructed by first PCR amplifying the *ori* and

kan^R genes from pUB110 using the primers KW513F and KW514R and then cloning the fragment into pUC19 using the SphI and HindIII restriction sites. The *G. thermoglucosidasius* 95A1 knockout plasmid pJZ04 was constructed by replacing the native *kan*^{R} gene in the plasmid pKW1012 with the *spc*^{R} gene from pDG1644. Briefly, the *spc*^{R} gene from pDG1644 was PCR amplified using the primers JZ019F and JZ020R and the pKW1012 plasmid was PCR amplified using the primers JZ017F and JZ018R. The two DNA fragments were then digested with SpeI and HindIII and ligated to obtain pJZ04. The *G. thermoglucosidasius* 95A1 knockout plasmid pJZ04s was constructed by removing the *amp*^{R} gene from the plasmid pJZ04. Briefly, the pJZ04 was PCR amplified using the primers of JZ310F and JZ311R. The fragment was then digested with EcoRI and self-ligated to obtain pJZ04s. The *G. thermoglucosidasius* C56-YS93 knockout plasmid pJZ04e was constructed by PCR amplifying the *mob* gene from pBBR1MCS-2 using the primers JZ102F and JZ103R and then cloning resulting DNA fragment into the plasmid pJZ04 using the NdeI and EcoRI restriction sites.

The *G. thermoglucosidasius* 95A1 *ldh* knockout plasmid pJZ16b (*spc*^{R}) was constructed as follows. The upstream fragment (600 bp) was PCR amplified from genomic DNA using the primers JZ055F and JZ056R and the downstream fragment (600 bp) was PCR amplified from genomic DNA using the primers JZ057F and JZ058R. The two DNA fragments were then ligated using NotI restriction site between them and cloned into the plasmid pJZ04 using the EcoRI and BamHI sites to obtain pJZ16b. The *G. thermoglucosidasius* C56-YS93 *ldh* knockout plasmid pJZ16c was constructed by cloning the 1.2 kb fragment from the plasmid pJZ16b using the EcoRI and BamHI restriction sites of pJZ04e. The *pfl* knockout plasmids pJZ07 (*spc*^{R}) for *G. thermoglucosidasius* 95A1 and pJZ07b (*spc*^{R}) for *G. thermoglucosidasius* C56-YS93 were constructed in similar manner using the primer pairs JZ025F/JZ026R to PCR amplify the

upstream region (600 bp) and JZ027F/JZ028R to amplify the downstream region (600 bp). The *G. thermoglucosidasius* 95A1 *aprt* and *spoIIIAA* knockout plasmids (spc^R) were constructed by using Gibson assembly [30]. The upstream (700 bp) and downstream (700 bp) regions of the *aprt* gene were PCR amplified using the respective primer pairs JZ334F/JZ335R and JZ336F/JZ337R. The plasmid pJZ04s was then digested with the restriction enzymes EcoRI and BamHI. The three fragments were then assembled using Gibson Assembly Cloning Kit (New England Biolabs) to obtain the plasmid pJZ58. The *spoIIIAA* knockout plasmid pJZ60 was similarly constructed using the primer pairs JZ353F/JZ354R and JZ355F/JZ356R to amplify the upstream (700 bp) and downstream (700 bp) regions, respectively.

The *pdc* expression plasmid pKW1024 ($chlor^R$) was constructed as follows. The *pdc* gene was PCR amplified from *S. venticuli* ATCC 55887 genomic DNA using the primers of KW519F2 and KW520R and the *recA* promoter was PCR amplified from *G. thermoglucosidasius* 95A1 genomic DNA using the primers of KW521F and KW522R. The two DNA fragments were then fused together using overlap PCR. The resulting fragment was then cloned into the XbaI and SphI sites of the plasmid pNW33N to obtain pKW1024.

3.5.3 Fermentations

Seed culture was prepared in 125 mL flasks with 20 mL TGP. The culture was inoculated with a single colony from a fresh plate and grown aerobically at 60°C, 250 rpm for 12 hour. Fermentation cultures were carried out in 15 mL sealed conical tubes containing 12 mL freshly prepared TMLM with glucose or cellobiose as carbon source. They were inoculated with the seed culture to an initial OD_{600} of 0.1 and grown at 60°C, 250 rpm for about 24 h. A separate tube (biological replicate) was used for each time-point measurement. The products were analyzed by HPLC.

Bioreactor experiments were carried out using a New Brunswick BioFlo/CelliGen 115 fermenter with 0.75 L working volume. The temperature was controlled at 60°C and the pH of the culture was maintained at 7.0 by addition of 4N NaOH or 4N H2SO4. Seed culture was prepared as above but using 500 mL flasks filled with 100 mL TGP. USYE was used as medium for the fermentation with 1% yeast extract and 3% glucose or 3% cellobiose as carbon source. Fermentation started with a 5% inoculation of seed culture. The bioreactor was operated with agitation of 400 rpm and aeration of 1 vvm at the beginning of fermentation. It was reduced to 200 rpm and 0.05 vvm when OD_{600} reached around 4. Antifoam 204 (Sigma) was added as required.

3.5.4 Strain adaptation

The double mutants of *G. thermoglucosidasius* (Δ*ldh* Δ*pfl*) were grown in 12 mL freshly prepared TMLM with 0.1% (v/v) acetic acid and 1% glucose or 1% cellobiose as carbon source in 15 mL sealed conical tubes until stationary phase. Cells were then serially transferred to fresh media for 10 times (4.17% inoculum for the first 5 transfers and then a 1.67% inoculum for the second 5 transfers), after which no further improvements in ethanol production were observed. Ten single colonies were isolated from the last subculture and the best ethanol producer among them was sent for genome sequencing. The adaptation experiments were performed in triplicate for each sugar.

3.5.5 Genome sequencing and analyses

Genomic DNA was prepared using the QIAamp DNA mini kit (Qiagen). Shotgun genome libraries were prepared and sequenced by the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign. The libraries were constructed using the Kapa DNA Library Construction kit (Kapa Biosystems) and quantitated by quantitative PCR. They were

then sequenced on a HiSeq2500 (Illumina) using a TruSeq SBS Sequencing Kit. Each sample yielded over 17 million reads. Fastq files were generated and de multiplexed using the bcl2fastq v1.8.4 Conversion Software (Illumina). Fastq files were analyzed using CLC Genomics Workbench version 6.5 (Qiagen). Reads were trimmed based on quality scores and then mapped to *G. thermoglucosidasius* 95A1 reference sequence.

3.6 Tables

Table 3.1 Fermentation profiles of engineered strains of *G. thermoglucosidasius* 95A1[a].

Strain	Residual glucose (g/L)	Ethanol (g/L)	Lactate (g/L)	Formate (g/L)	Acetate[b] (g/L)	Ethanol yield (g/g)	Carbon recov. (%)
Wild type	2.71±0.44	0.54±0.04	4.20±0.51	0.31±0.14	0.47±0.03	0.06±0.01	8.30±0.85
Δldh	2.32±0.37	2.01±0.17	0.08±0.01	1.20±0.02	0.93±0.01	0.24±0.02	31.47±2.25
PrecA::pdc Δldh	2.50±1.13	2.34±0.45	0.08±0.01	1.20±0.08	0.90±0.05	0.26±0.02	33.94±2.74
Δldh Δpfl[c]	7.48±0.66	1.24±0.09	-	-	0.73±0.01	0.38±0.08	43.09±0.09
Evolved strain (JZ03d1)[c]	0.26±0.03	4.25±0.15	-	-	1.16±0.02	0.43±0.02	55.59±2.53
Δldh Δpfl Δaprt[c]	2.49±0.55	3.46±0.15	-	-	0.69±0.27	0.42±0.01	51.72±2.41
Δldh Δpfl ΔspoIIIAA[c]	4.98±0.23	2.16±0.02	-	-	1.03±0.01	0.37±0.04	47.08±4.36
Δldh Δpfl Δaprt ΔspoIIIAA[c]	1.63±0.21	3.69±0.06	-	-	0.64±0.04	0.40±0.01	48.58±1.48

[a]Cells were cultured in TMLM with 1% glucose.

[b]Acetate in the supernatant.

[c]Medium was supplemented with 0.1% acetic acid.

Table 3.2 Fermentation profiles of engineered strains of *G. thermoglucosidasius* 95A1[a].

Strain	Residual cellobiose (g/L)	Ethanol (g/L)	Lactate (g/L)	Formate (g/L)	Acetate[b] (g/L)	Ethanol yield (g/g)	Carbon recov. (%)
Wild type	2.65±0.64	0.51±0.07	4.80±0.48	0.33±0.05	0.58±0.05	0.07±0.01	8.58±0.91
Δ*ldh*	2.82±0.58	1.73±0.16	0.02±0.04	1.29±0.14	1.13±0.12	0.24±0.02	29.86±2.22
P*recA*::*pdc* Δ*ldh*	2.62±0.41	2.17±0.18	0.05±0.04	1.59±0.06	1.23±0.04	0.24±0.03	29.99±4.00
Δ*ldh* Δ*pfl*[b]	6.57±0.93	1.36±0.47	-	-	1.11±0.12	0.41±0.01	47.98±0.34
Evolved strain (JZ03c1)	1.83±1.99	3.62±0.56	-	-	1.29±0.06	0.46±0.03	57.50±4.17
Δ*ldh* Δ*pfl* Δ*aprt*	2.83±0.19	2.86±0.02	-	-	1.06±0.29	0.42±0.01	49.87±1.93
Δ*ldh* Δ*pfl* Δ*spoIIIAA*	5.37±0.05	1.69±0.01	-	-	1.29±0.05	0.32±0.04	39.42±6.03
Δ*ldh* Δ*pfl* Δ*aprt* Δ*spoIIIAA*	2.86±1.28	2.99±0.44	-	-	1.18±0.06	0.45±0.02	54.44±2.47

[a]Cells were cultured in TMLM with 1% cellobiose.

[b]Acetate in the supernatant.

[c]Medium was supplemented with 0.1% acetic acid for this strain and others below.

Table 3.3 Single nucleotide polymorphisms (SNPs) identified in the 95A1 strains evolved in glucose (JZ03d1, JZ03d2 and JZ03d3) or cellobiose (JZ03c1, JZ03c2 and JZ03c3)

Non-synonymous SNPs in CDS	**JZ03d1**	**JZ03d2**	**JZ03d3**	**JZ03c1**	**JZ03c2**	**JZ03c3**
aprt	Phe74Val	Thr123Ile	Phe74Val	Frame shift after Cys64	Thr123Ile	Frame shift after Cys64
spoIIIAA	Frame shift after Ile4	Frame shift after Ile4	Frame shift after Ile4			
spoVAD				Ala170Val		
yabK				Glu56Lys		Glu56Lys
duf881					Met1Arg	
mazG						Glu224Lys
ald						Thr194Met
Total SNPs in CDS	3	4	3	4	2	5
Total SNPs	4	5	4	9	3	8

Table 3.4 Identified SNPs unique to evolved strains in glucose (JZ03d1, JZ03d2, JZ03d3) or cellobiose (JZ03c1, JZ03c2, JZ03c3) relative to the parent strain 95A1 Δ*ldh* Δ*pfl* (JZ03)

Strains	Reference Position	Annotations	Coding Region Change	Amino Acid Change
JZ03d1	295821	*aprt*	220T>G	Phe74Val
	451621^451622	*spoIIIAA*	13_14insG	Frame shift after Ile4
	489200	Intergenic		
	1282791	GY4MC1_2354	117C>T	No AA change
JZ03d2	295969	*aprt*	368C>T	Thr123Ile
	451621^451622	*spoIIIAA*	13_14insG	Frame shift after Ile4
	489200	Intergenic		
	1066527	*cbiA*	225T>C	No AA change
	1282791	GY4MC1_2354	117C>T	No AA change
JZ03d3	295821	*aprt*	220T>G	Phe74Val
	451621^451622	*spoIIIAA*	13_14insG	Frame shift after Ile4
	489200	Intergenic		
	1282791	GY4MC1_2354	117C>T	No AA change
JZ03c1	295792^295793	*aprt*	191_192insC	Frame shift after Cys64
	779378	Intergenic		
	1282791	GY4MC1_2354	117C>T	No AA change
	2296935	*spoVAD*	509C>T	Ala170Val
	2569248	Intergenic		
	2569250	Intergenic		
	2569253	Intergenic		
	3132445	Intergenic		
	3162036	*yabK*	166G>A	Glu56Lys
JZ03c2	295969	*aprt*	368C>T	Thr123Ile
	2083133	*duf881*	2T>G	Met1Arg
	3132445	Intergenic		
JZ03c3	295792^295793	*aprt*	191_192insC	Frame shift after Cys64
	489200	Intergenic		

Table 3.4 (cont.)

Strains	Reference Position	Annotations	Coding Region Change	Amino Acid Change
	1282791	GY4MC1_2354	117C>T	No AA change
	3132445	Intergenic		
	3155014	*mazG*	670G>A	Glu>Lys
	3162036	*yabK*	166G>A	Glu56Lys
	3287557	*ald*	581C>T	Thr>Met
	3848260	Geoth_R0098	52C>T	

Table 3.5 Fermentation profiles of engineered strains of *G. thermoglucosidasius* C56-YS93[a].

Strain	Residual glucose (g/L)	Ethanol (g/L)	Lactate (g/L)	Formate (g/L)	Acetate[b] (g/L)	Ethanol yield (g/g)	Carbon recovery
Wild type	4.76±0.01	0.17±0.01	4.86±0.01	1.17±0.01	0.50±0.01	0.02±0.01	2.61±0.15%
Δ*ldh*	8.72±0.42	0.15±0.03	0.07±0.01	1.07±0.14	0.78±0.08	0.05±0.01	6.49±0.40%
Δ*ldh* Δ*pfl*	8.90±0.23	0.07±0.01	-	-	1.12±0.01	0.03±0.01	3.83±0.18%
Evolved strain[c]	8.23±0.24	0.21±0.01	-	-	1.02±0.02	0.06±0.01	7.51±0.14%

[a]Cells were cultured in TMLM with 1% glucose.

[b]Acetate in the supernatant.

[c]Medium was supplemented with 0.1% acetic acid.

Table 3.6 Strains and plasmids used in this study

Plasmid or strain	Characteristics	Source or reference
Plasmids		
pBBR1MCS-2	Source of *mob*	[31]
pDG1664	Source of *spc*	BGSC
pNW33N	Cm^R; *E. coli-Bacillus* shuttle vector	BGSC
pKW1012	Kan^R, Amp^R; *E. coli-Bacillus* shuttle vector containing *repB* and Kan^R from pUB110 in mcs of pUC19	This study
pKW1024	pNW33N containing *pdc* (AF354297.1) from *Sarcina venticuli* with *recA* promoter from *G. thermoglucosidasius* 95A1	This study
pJZ04	Spc^R replaces of Kan^R in pKW1012	This study
pJZ04e	pJZ04 containing *mob* from pBBR1MCS-3	This study
pJZ07	pJZ04 containing truncated *pfl* (Geoth_3895)	This study
pJZ07b	pJZ04e containing truncated *pfl*	This study
pJZ16b	pJZ04 containing truncated *ldh* (Geoth_3351)	This study
pJZ16c	pJZ04e containing truncated *ldh*	This study
pJZ58	pJZ04 containing truncated *aprt* (Geoth_1059)	This study
pJZ60	pJZ04 containing truncated *spoIIIAA* (Geoth_1270)	This study
Strains		
E. coli BW20767	*RP4-2*(Tet::*Mu-1*, Kan::*Tn7*) Δ *uidA*::*pir*$^+$ *recA1 creB510 leu-63 hsdR17 endA1 zbf-5 thi*	ATCC47084
G. thermoglucosidasius BGSC 95A1	Wild type, $DSM2452^T$	BGSC
G. thermoglucosidasius C56-YS93	Wild type	
JZ01	Δ*ldh* variant of *G. thermoglucosidasius* 95A1	This study
JZ02	P*recA*::*pdc* Δ*ldh* variant of *G. thermoglucosidasius* 95A1	This study
JZ03	Δ*ldh* Δ*pfl* variant of *G. thermoglucosidasius* 95A1	This study
JZ03c	Evolved strains of JZ03 in cellobiose (JZ03c1, JZ03c2 and JZ03c3)	This study
JZ03d	Evolved strains of JZ03 in glucose (JZ03d1, JZ03d2 and JZ03d3)	This study
JZ04	Δ*ldh* variant of *G. thermoglucosidasius* C56-YS93	This study
JZ05	Δ*ldh* Δ*pfl* variant of *G. thermoglucosidasius* C56-YS93	This study

Table 3.6 (cont.)

Plasmid or strain	Characteristics	Source or reference
JZ05d	Evolved strains of JZ05 in glucose	This study
JZ06	Δ*ldh* Δ*pfl* Δ*aprt* variant of *G. thermoglucosidasius* 95A1	This study
JZ07	Δ*ldh* Δ*pfl*Δ*spoIIIAA* variant of *G. thermoglucosidasius* 95A1	This study
JZ08	Δ*ldh* Δ*pfl* Δ*aprt* Δ*spoIIIAA* variant of *G. thermoglucosidasius* 95A1	This study

Table 3.7 Primers used in this study

Primer	Sequence	Plasmid
KW519F2	ATGAAAATAACAATTGCAGAATACTT	pKW1024
KW520R	ATAGCATGCTTAGTAGTTATTTTGAGAACTAAATAGACTTGC	
KW521F	ATATCTAGACACGTTCCCGCTTCTGTTTTAGCA	
KW522R	AAGTATTCTGCAATTGTTATTTTCATTAACAAACTCCTCCTTTTTCTTGAC	
JZ017F	CCCAAGCTTGGCGTAATCATGGTCATAGCTGTTTC	pJZ04
JZ018R	GGACTAGTTCCCTTTTCAGATAATTTTAGATTTGC	
JZ019F	GGACTAGTATGAGCAATTTGATTAACGG	
JZ020R	CCCAAGCTTCTAATTGAGAGAAGTTTCTATAG	
JZ025F	ATAGAATTCGGCTTGCCATTTCTTCTTC	pJZ07
JZ026R	ATAGCGGCCGCGTGTGTTTTCCGGTATTCC	
JZ027F	ATAGCGGCCGCGAAATGGAAGGCGACTTC	
JZ028R	CGGGATCCTTACATCGTTTCGTGGAACGTG	
JZ055F	CGGAATTCGTCATTGGCGACGATCAGCTTC	pJZ16b
JZ056R	ATAGCGGCCGCCGCTGTCTGTCATCCTTTC	
JZ057F	ATAGCGGCCGCTGATGCGGCATATCAAATCATTG	
JZ058R	CGGGATCCTATACAGCCAGACGGCCATAATC	
JZ102F	ATAAAGCTTTTGATGGGGCAAGGCCGCAGGCCG	pJZ04e
JZ103R	AACTGCAGGCATGCTCATGATAATAATGGTTTCTTAG	
JZ310F	TCAGAATTCGAGCTCGGTACCCGGGGATCC	pJZ04s
JZ311R	CGGAATTCTAACTGTCAGACCAAGTTTAC	
JZ334F	TGAGTAAACTTGGTCTGACAGTTAGAATTCTATGATATGGCGGTGCATGAATG	pJZ58
JZ335R	AAAGGCAATGCCGGCGACGACGCCGAATCTGGTCCGTGGCATATTTATAC	
JZ336F	CGGCGTCGTCGCCGGCATTGCCTTTTTG	
JZ337R	GCATGCCTGCAGGTCGACTCTAGAGGATCCAAAATCTCCAACGTTTCATTG	
JZ353F	TGAGTAAACTTGGTCTGACAGTTAGAATTCGCCCCTTGCGCTTGGCGCGG	pJZ60
JZ354R	ACCGTTGTCCATATGCATACTCCGGTGGAAGAAATTGAGCCGTTCCTTTC	
JZ355F	CCGGAGTATGCATATGGACAAC	
JZ356R	GCATGCCTGCAGGTCGACTCTAGAGGATCCTAAAATCACCAGAAGCAGTC	

3.7 Figures

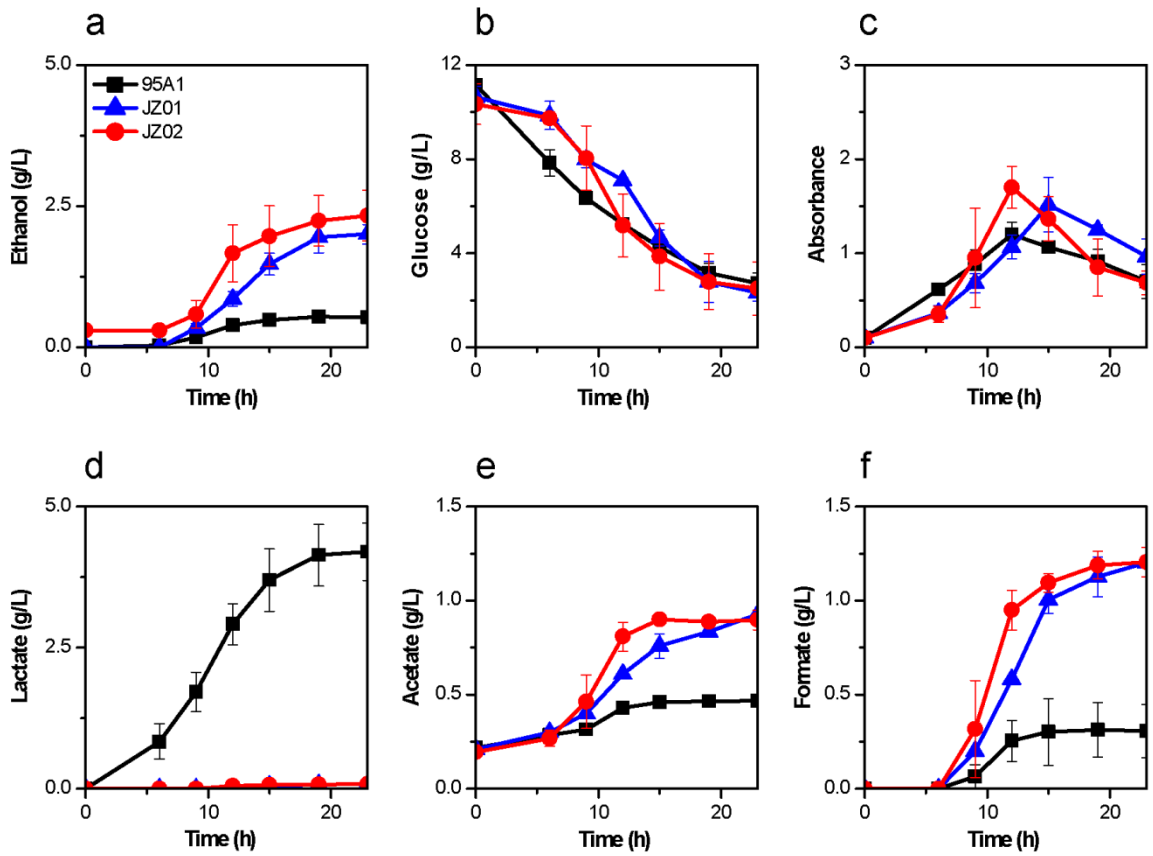

Figure 3.1 Ethanol production in engineered *G. thermoglucosidasius* 95A1. Cells were cultured in TMLM with 1% glucose. (a-b) Ethanol and glucose concentrations in *G. thermoglucosidasius* 95A1, Δ*ldh* (JZ01) and P*recA*::*pdc* Δ*ldh* (JZ02). (c) Cell growth of the strains as determined by OD_{600}. (d-f) Lactate, acetate and formate concentrations. The figures show the means and standard deviations obtained from three experiments performed on separate days.

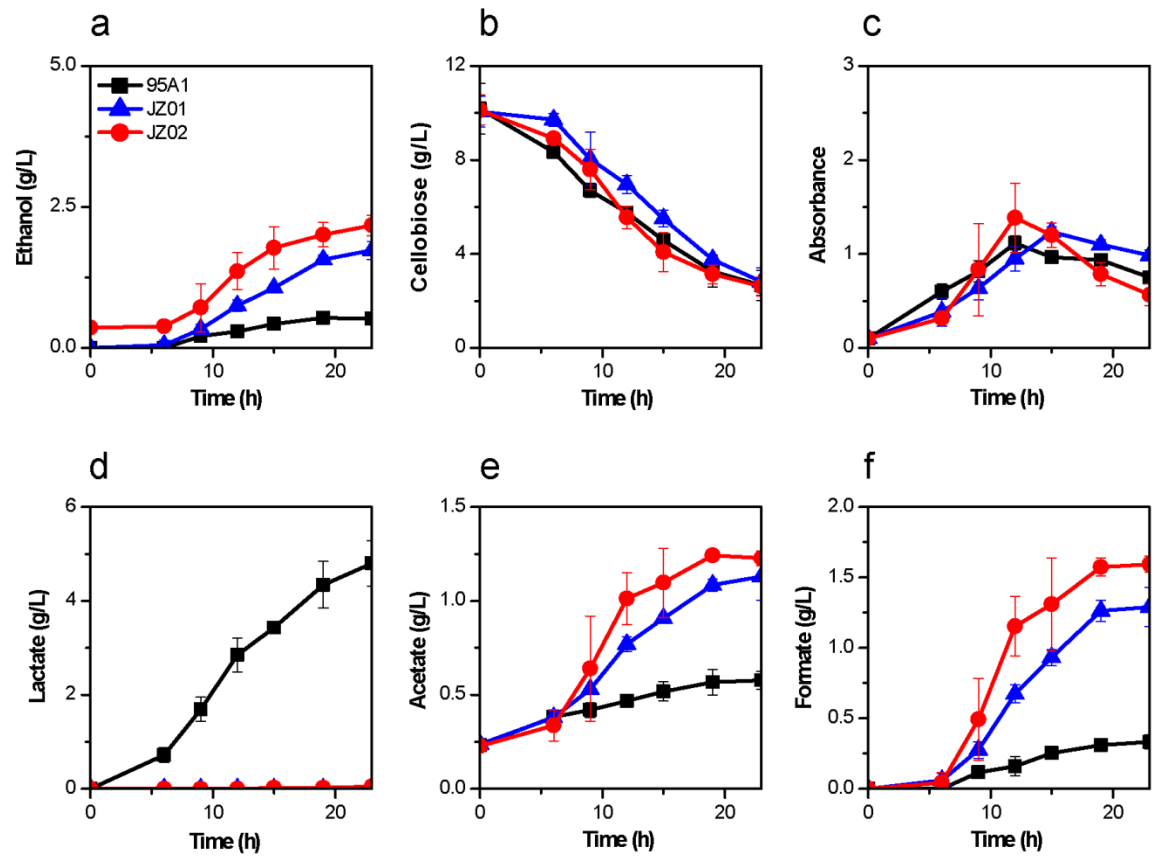

Figure 3.2 Ethanol production in engineered *G. thermoglucosidasius* 95A1. Cells were cultured in TMLM with 1% cellobiose. (a-b) Ethanol and cellobiose concentrations of *G. thermoglucosidasius* 95A1, Δ*ldh* (JZ01) and P*recA*::*pdc* Δ*ldh* (JZ02). (c) Cell growth of the strains as determined by OD_{600}. (d-f) Lactate, acetate and formate concentrations. The figures show the means and standard deviations obtained from three experiments performed on separate days.

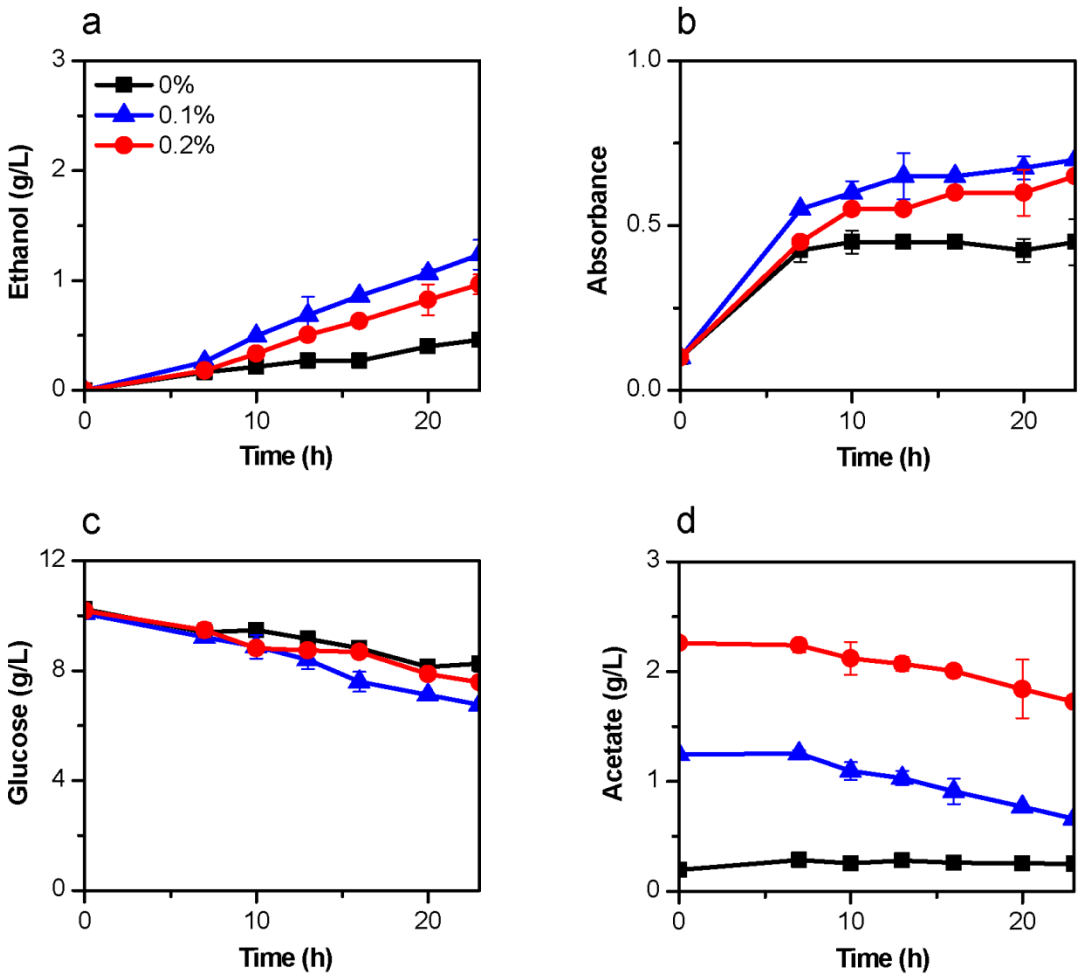

Figure 3.3 Effect of acetic acid on ethanol production in 95A1 Δ*ldh* Δ*pfl* (JZ03). Cells were cultured in TMLM with 1% glucose and acetic acid at different concentrations. (a) Ethanol concentrations of the strain at different concentrations of acetic acid. (b) Cell growth as determined by OD_{600}. (c-d) Glucose and acetate concentrations. The figures show the means and standard deviations obtained from two experiments.

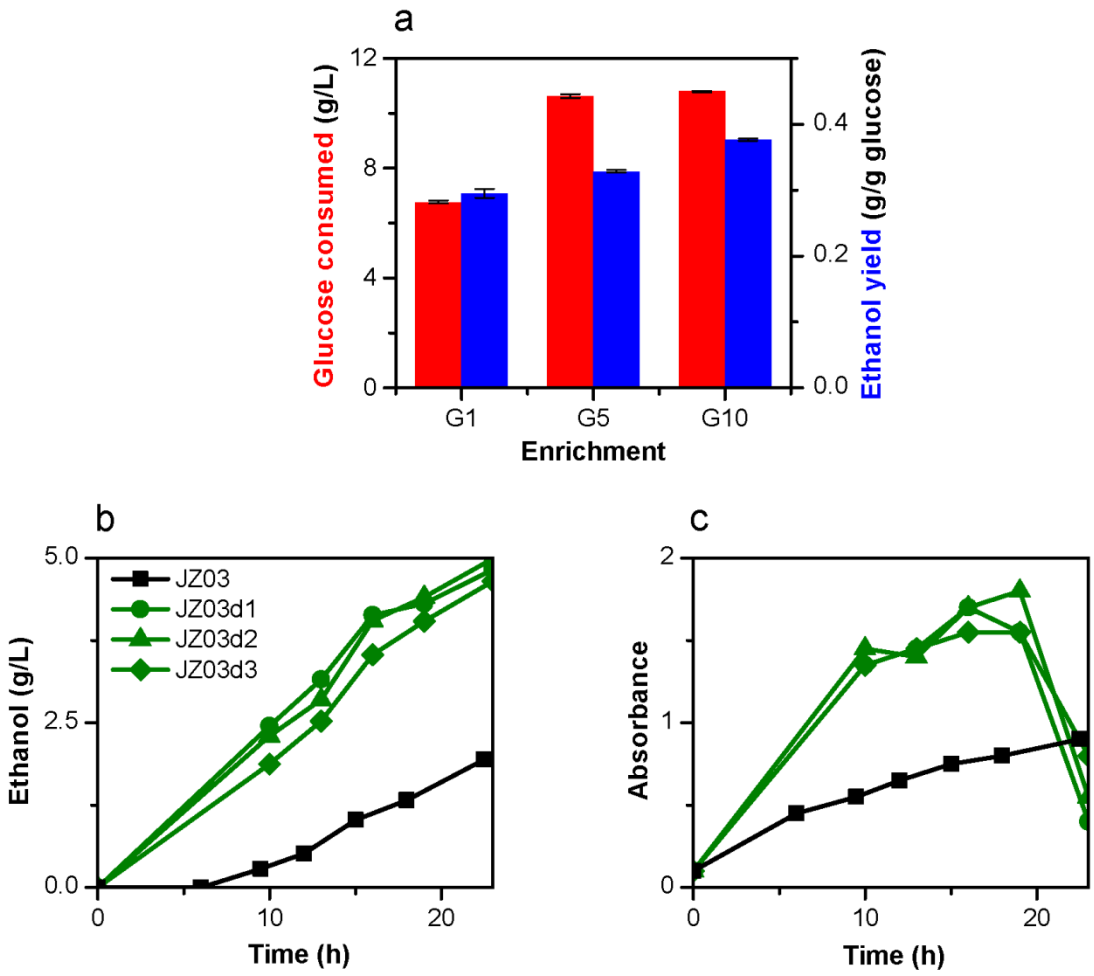

Figure 3.4 Evolution of the strain 95A1 Δ*ldh* Δ*pfl* (JZ03) in glucose. (a) Glucose consumption and ethanol yield of the strain during serial transfer in TMLM with 1% glucose and 0.1% acetic acid. The figure illustrates the means and standard deviations of duplicate experiments. (b) Ethanol concentrations of the parent strain (JZ03) and three evolved strains (JZ03d1, JZ03d2 and JZ03d3) from the last subculture. (c) Cell growth of the strains as determined by OD_{600}.

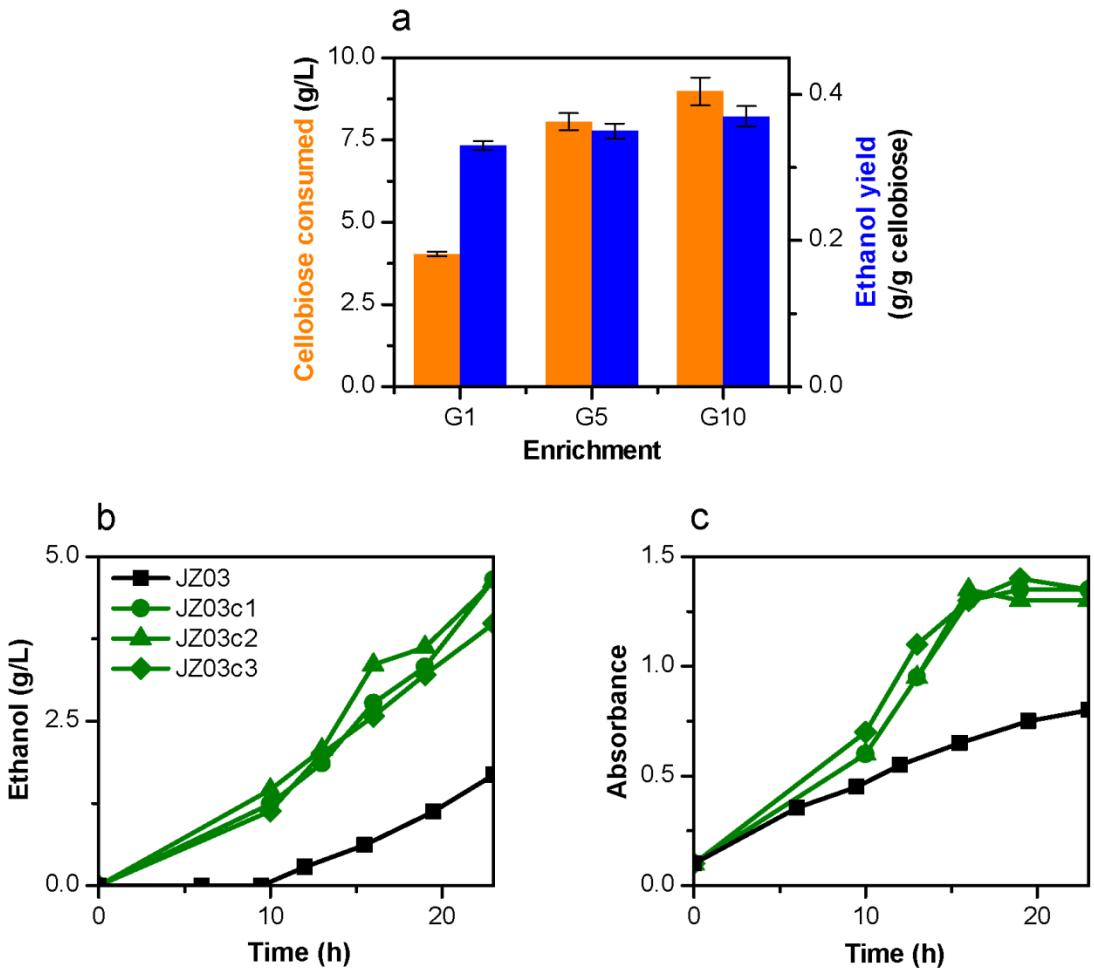

Figure 3.5 Evolution of the strain 95A1 Δ*ldh* Δ*pfl* (JZ03) in cellobiose. (a) Cellobiose consumption and ethanol yield of the strain during serial transfer in TMLM with 1% cellobiose and 0.1% acetic acid. The figures show the means and standard deviations obtained from two experiments. (b) Ethanol concentrations of the parent strain (JZ03) and three evolved strains (JZ03c1, JZ03c2 and JZ03c3) from the last subculture. (c) Cell growth of the strains as determined by OD_{600}.

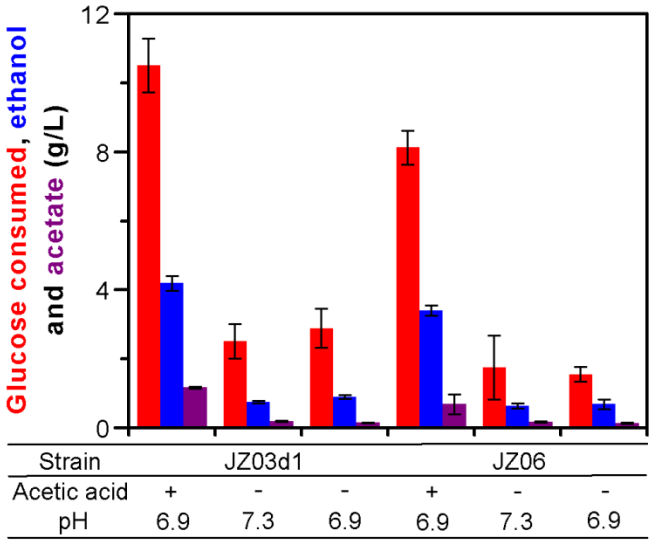

Figure 3.6 Effect of acetic acid on ethanol production in the evolved strain (JZ03d1) and Δ*ldh* Δ*pfl* Δ*aprt* (JZ06). Cells were cultured in TMLM with 1% glucose in the absence or presence of 0.1% acetic acid (indicated by +). The supplementation of acetic acid decreased the pH of the starter culture from 7.3 to 6.9. A control was made by adjusting the pH to 6.9 using 1 M HCl. The figures show the means and standard deviations obtained from three experiments.

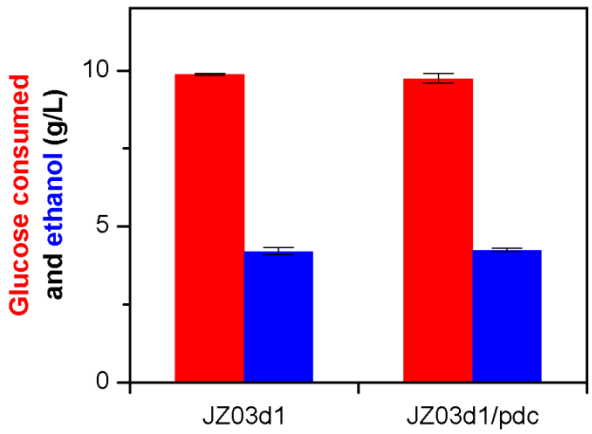

Figure 3.7 Pdc expression in the evolved strain (JZ03d1). Cells were cultured in TMLM with 1% glucose and 0.1% acetic acid. The figure shows the means and standard deviations obtained from three experiments.

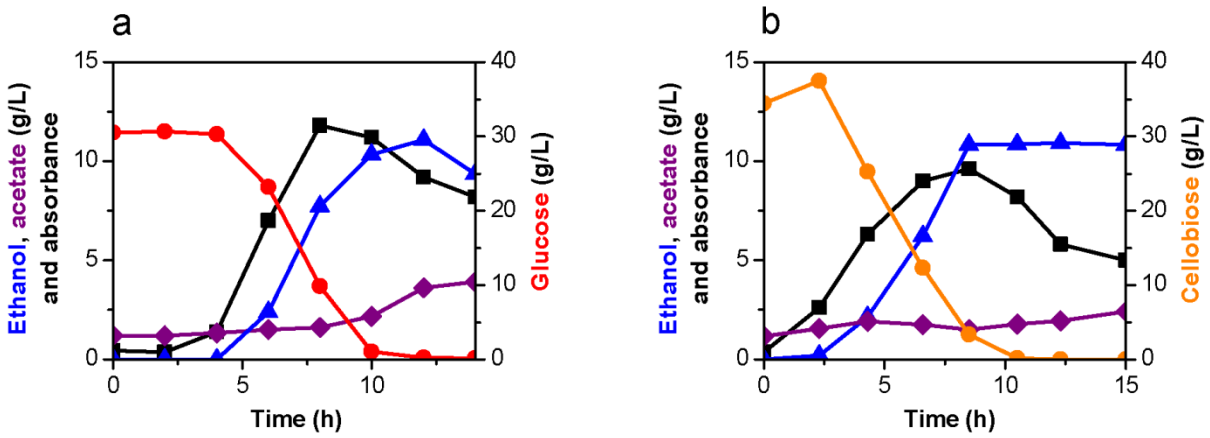

Figure 3.8 Batch growths of evolved strains in bioreactors. . Minimal medium USYE was used with 0.1% acetic acid and 3% glucose or 3% cellobiose. (a) Glucose-evolved strain (JZ03d1). (b) Cellobiose-evolved strain (JZ03c1).

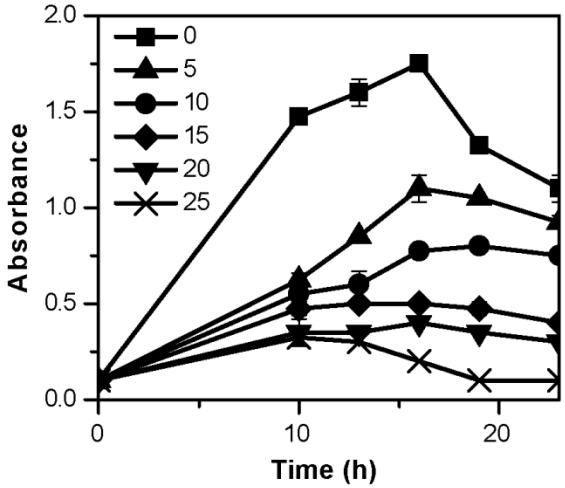

Figure 3.9 Ethanol tolerance of the evolved strain (JZ03d1). Cells were cultured in TMLM with 1% glucose and different concentrations of ethanol from 0 to 25 g/L. The figures show the means and standard deviations obtained from two experiments.

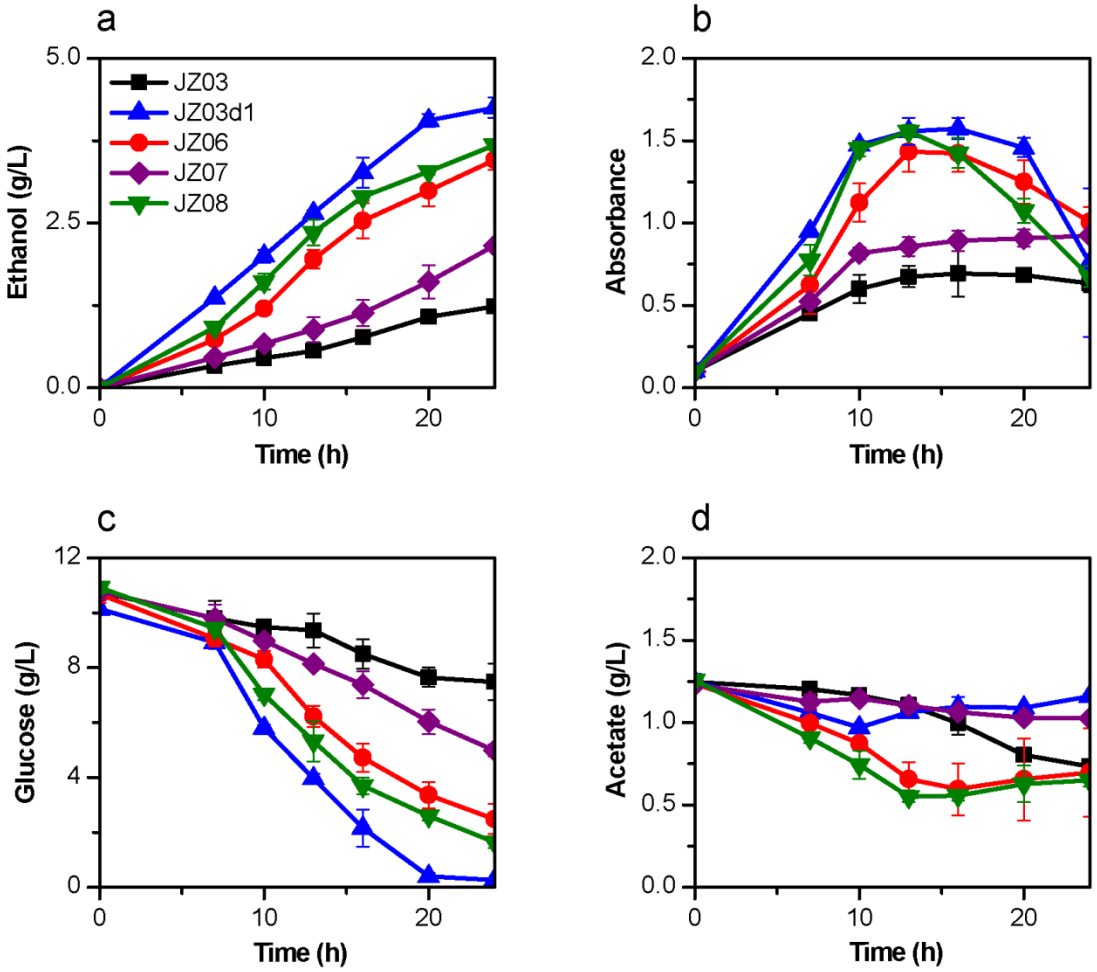

Figure 3.10 Effect of *aprt* and *spoIIIAA* mutations on ethanol production. Cells were cultured in TMLM with 1% glucose and 0.1% acetic acid. (a) Ethanol concentrations of the parent strain Δ*ldh* Δ*pfl* (JZ03), glucose-evolved strain (JZ03d1), Δ*ldh* Δ*pfl* Δ*aprt* (JZ06), Δ*ldh* Δ*pfl* Δ*spoIIIAA* (JZ07), Δ*ldh* Δ*pfl* Δ*aprt* Δ*spoIIIAA* (JZ08). (b) Cell growth of the strains determined by OD_{600}. (c-d) Glucose and acetate concentrations. The figures show the means and standard deviations obtained from three experiments performed on separate days.

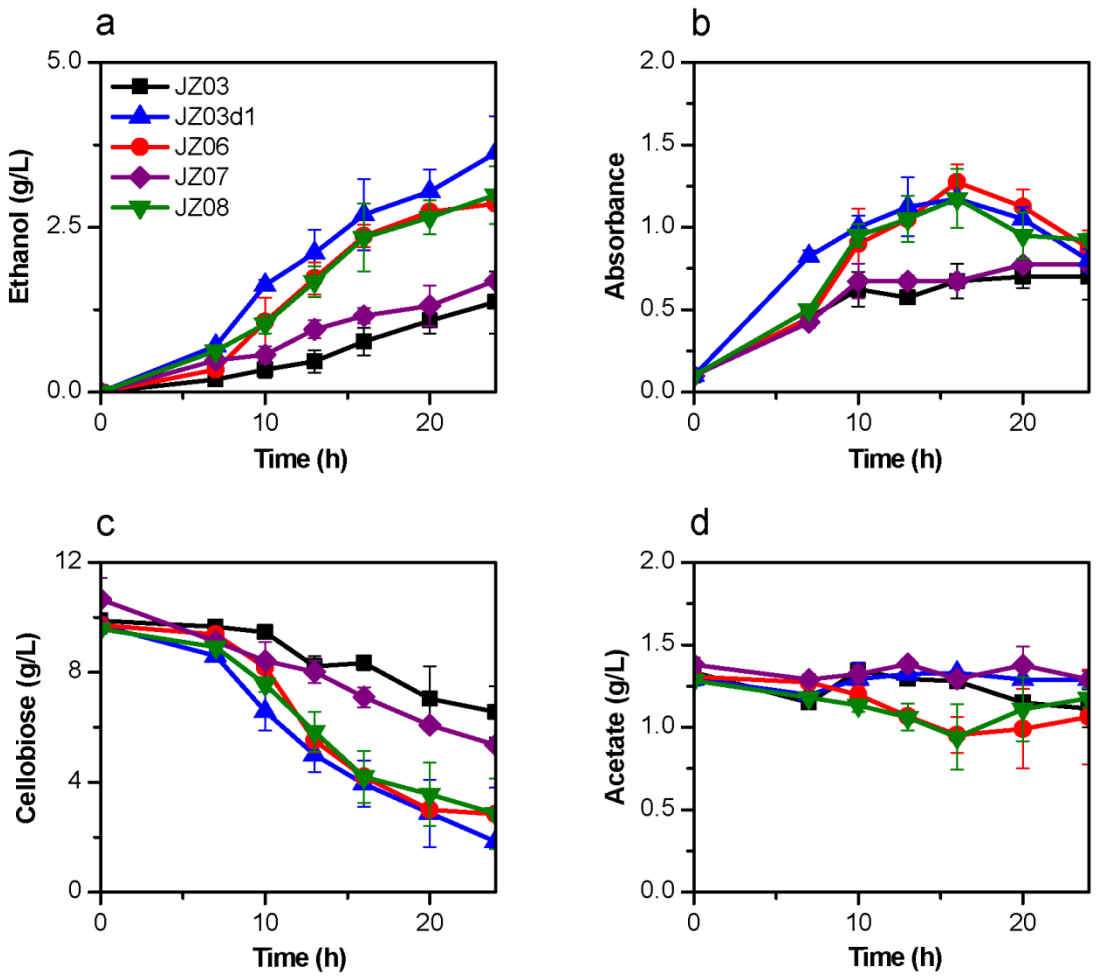

Figure 3.11 Effect of *aprt* and *spoIIIAA* mutations on ethanol production. Cells were cultured in TMLM with 1% cellobiose and 0.1% acetic acid. (a) Ethanol concentrations ofthe parent strain Δ*ldh* Δ*pfl* (JZ03), cellobiose-evolved strain (JZ03c1), Δ*ldh* Δ*pfl* Δ*aprt* (JZ06), Δ*ldh* Δ*pfl* Δ*spoIIIAA* (JZ07), Δ*ldh* Δ*pfl* Δ*aprt* Δ*spoIIIAA* (JZ08). (b) Cell growth of the strains determined by OD_{600}. (c-d) Cellobiose and acetate concentrations. The figures show the means and standard deviations obtained from two experiments.

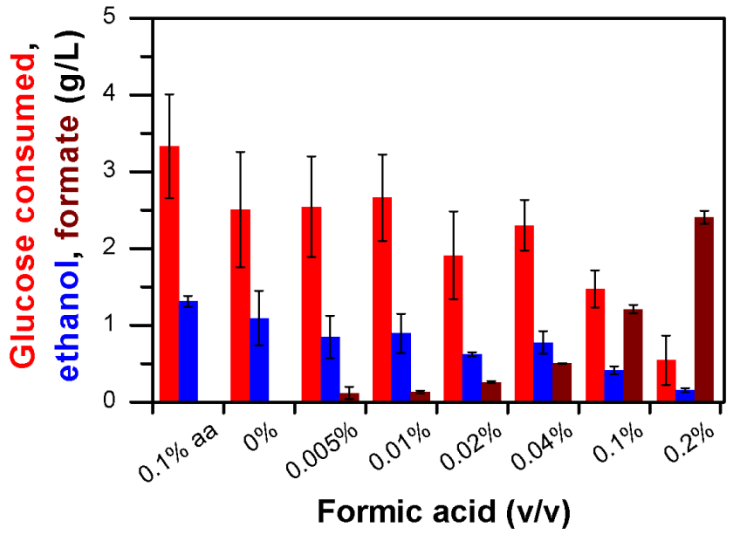

Figure 3.12 Effect of formic acid on ethanol production in 95A1 Δ*ldh* Δ*pfl* (JZ03). Cells were cultured in TMLM with 1% glucose and formic acid at different concentrations. They were compared with the one supplemented with 0.1% acetic acid (indicated by 0.1% aa). The figures show the means and standard deviations obtained from three experiments.

3.8 References

1. Chung D, Cha M, Guss AM, Westpheling J (2014) Direct conversion of plant biomass to ethanol by engineered Caldicellulosiruptor bescii. Proceedings of the National Academy of Sciences 111: 8931-8936.
2. Herring CD, Kenealy WR, Shaw AJ, Raman B, Tschaplinski TJ, et al. (2012) Final Report on Development of Thermoanaerobacterium saccharolyticum for the conversion of lignocellulose to ethanol. DOE/GO/017057-1; TRN: US201212%%909 United States10.2172/1033560TRN: US201212%%909Wed Dec 05 11:12:50 EST 2012GFOEnglish DOE/GO/017057-1; TRN: US201212%%909 United States10.2172/1033560TRN: US201212%%909Wed Dec 05 11:12:50 EST 2012GFOEnglish. Medium: ED; Size: 106 Kb p.
3. Shaw AJ, Miller BB, Rogers SR, Kenealy WR, Meola A, et al. (2015) Anaerobic detoxification of acetic acid in a thermophilic ethanologen. Biotechnology for Biofuels 8.
4. Cripps RE, Eley K, Leak DJ, Rudd B, Taylor M, et al. (2009) Metabolic engineering of Geobacillus thermoglucosidasius for high yield ethanol production. Metabolic Engineering 11: 398-408.
5. Taylor MP, Esteban CD, Leak DJ (2008) Development of a versatile shuttle vector for gene expression in Geobacillus spp. Plasmid 60: 45-52.
6. Taylor MP, van Zyl L, Tuffin IM, Leak DJ, Cowan DA (2011) Genetic tool development underpins recent advances in thermophilic whole-cell biocatalysts. Microbial Biotechnology 4: 438-448.
7. Van Zyl LJ, Taylor MP, Eley K, Tuffin M, Cowan DA (2013) Engineering pyruvate decarboxylase-mediated ethanol production in the thermophilic host Geobacillus thermoglucosidasius. Applied Microbiology and Biotechnology 98: 1247-1259.
8. Zhou J, Wu K, Rao CV (2016) Evolutionary engineering of Geobacillus thermoglucosidasius for improved ethanol production. Biotechnology and Bioengineering 113: 2156-2167.
9. Talarico LA (2005) Construction and expression of an ethanol production operon in Gram-positive bacteria. Microbiology 151: 4023-4031.
10. San Martin R, Bushell D, Leak DJ, Hartley BS (1994) Cultivation of an LL-lactate dehydrogenase mutant of Bacillus stearothermophilus in continuousculture with cell recycle. Biotechnology and Bioengineering 44: 21–28.
11. Thompson AH, Studholme DJ, Green EM, Leak DJ (2008) Heterologous expression of pyruvate decarboxylase in Geobacillus thermoglucosidasius. Biotechnology Letters 30: 1359-1365.
12. Starai VJ, Escalante-Semerena JC (2004) Acetyl-coenzyme A synthetase (AMP forming). Cellular and Molecular Life Sciences 61.
13. Gardner JG, Grundy FJ, Henkin TM, Escalante-Semerena JC (2006) Control of Acetyl-Coenzyme A Synthetase (AcsA) Activity by Acetylation/Deacetylation without NAD+ Involvement in Bacillus subtilis. Journal of Bacteriology 188: 5460-5468.
14. Hochstadt-Ozer J (1972) The regulation of purine utilization in bacteria. J Biol Chem 247: 2419-2426.
15. Copenhaver GP, Doan T, Morlot C, Meisner J, Serrano M, et al. (2009) Novel Secretion Apparatus Maintains Spore Integrity and Developmental Gene Expression in Bacillus subtilis. PLoS Genetics 5: e1000566.
16. Argyros DA, Tripathi SA, Barrett TF, Rogers SR, Feinberg LF, et al. (2011) High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. Applied and Environmental Microbiology 77: 8288-8294.
17. Basen M, Schut GJ, Nguyen DM, Lipscomb GL, Benn RA, et al. (2014) Single gene insertion drives bioalcohol production by a thermophilic archaeon. Proceedings of the National Academy of Sciences 111: 17618-17623.
18. Lavania M, Cheema S, Sarma PM, Ganapathi R, Lal B (2014) Methanogenic potential of a thermophilic consortium enriched from coal mine. International Biodeterioration & Biodegradation 93: 177-185.
19. Lin PP, Rabe KS, Takasumi JL, Kadisch M, Arnold FH, et al. (2014) Isobutanol production at elevated temperatures in thermophilic Geobacillus thermoglucosidasius. Metabolic Engineering 24: 1-8.
20. Talluri S, Raj SM, Christopher LP (2013) Consolidated bioprocessing of untreated switchgrass to hydrogen by the extreme thermophile Caldicellulosiruptor saccharolyticus DSM 8903. Bioresource technology 139: 272-279.
21. Wolfe AJ (2005) The acetate switch. Microbiology and molecular biology reviews 69: 12-50.

22. Palmqvist E, Hahn-Hägerdal B (2000) Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. Bioresource technology 74: 25-33.
23. Wei N, Quarterman J, Kim SR, Cate JHD, Jin Y-S (2013) Enhanced biofuel production through coupled acetic acid and xylose consumption by engineered yeast. Nature Communications 4.
24. Leibig M, Liebeke M, Mader D, Lalk M, Peschel A, et al. (2011) Pyruvate formate lyase acts as a formate supplier for metabolic processes during anaerobiosis in Staphylococcus aureus. Journal of Bacteriology 193: 952-962.
25. Rydzak T, Lynd LR, Guss AM (2015) Elimination of formate production in Clostridium thermocellum. Journal of industrial microbiology & biotechnology 42: 1263-1272.
26. Suzuki H (2012) Genetic Transformation of Geobacillus kaustophilus HTA426 by Conjugative Transfer of Host-Mimicking Plasmids. Journal of Microbiology and Biotechnology 22: 1279-1287.
27. Kananavičiūtė R, Čitavičius D (2015) Genetic engineering of Geobacillus spp. Journal of Microbiological Methods 111: 31-39.
28. Liao HH, Kanikula AM (1990) Increased efficiency of transformation ofBacillus stearothermophilus by a plasmid carrying a thermostable kanamycin resistance marker. Current microbiology 21: 301-306.
29. Guérout-Fleury A-M, Frandsen N, Stragier P (1996) Plasmids for ectopic integration in Bacillus subtilis. Gene 180: 57-61.
30. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature methods 6: 343-345.
31. Kovach M, Elzer P, Hill D, Robertson G, Farris M, et al. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166: 175-176.

Chapter 4 Engineer *G. thermoglucosidasius* for *2R, 3R*-Butanediol (*R*-BDO)

4.1 Introduction

2, 3-Butanediol (BDO) has attracted growing attention due to its important industrial applications [1,2]. It is used as an antifreeze agent because of its low freezing point. It is also an important chemical intermediate used for the production of synthetic rubber, fuel and drugs [1]. BDO has three isomeric forms existing in nature, (*2S*, *3S*)-butanediol (*S*-BDO), *meso*-butanediol (*meso*-BDO), and (*2R*, *3R*)-butanediol (*R*-BDO). A lot of microorganisms can natively produce BDO, especially *Klebsiella* and *Enterobacter* species, which can accumulate BDO to high levels. However, these native producers are not commercial friendly enough, since they are pathogenic or the BDO synthesized is generally a mixture of stereoisomers [3]. Therefore, some industrially friendly hosts are being developed for BDO production, such as *E. coli* [3], *B. subtilis* [4] and *S. cerevisiae* [5,6].

According to genomic analysis, *G. thermoglucosidasius* only has the pathways to produce *R*-BDO (Figure 4.1). The pathways include four reactions, which involve a spontaneous step requiring O_2. The other three reactions are catalyzed by two enzymes, acetolactate synthase (*alsS*) and *2R*, *3R*-butanediol dehydrogenase (*bdhA*). In this Chapter, *G. thermoglucosidasius* 95A1was engineered for improved enantiopure *R*-BDO production. To eliminate the requirement of O_2, a new pathway to synthesize BDO was introduced by expressing a thermostable heterologous acetolactate decarboxylase (*alsD*). Since the native AlsS was reported to be inefficient [7], thermostable heterologous AlsS was co-expressed. In order to optimize the new pathway, different potentially strong promoters were screened. For the best combination of enzymes for *R*-

BDO production, different *alsS* and *alsD* genes from thermophiles were cloned and tested. As a result, the strain co-expressing acetolactate synthase from *B. subtilis* and acetolactate decarboxylase from *Streptococcus thermophilus* under the lactate dehydrogenase promoter from *G. thermodenitrificans*, had the highest *R*-BDO production. The fermentation conditions were optimized to increase the production furtherly. Different alcohol dehydrogenase (*adh*) were deleted separately to divert more carbon source to *R*-BDO, but it only increased the yield but not production. The production of *R*-BDO was achieved in this chapter, which was aimed to confirm *G. thermoglucosidasius* as a potential candidate for valuable biochemical production.

4.2 Results

4.2.1 Genetic tools development for pathway construction

In our experience of transforming *G. thermoglucosidasius* 95A1 by electroporation, the efficiency decreased as the size of plasmid became bigger. Thus, the shuttle vector, pNW33N, usually used for protein expression in *Geobacillus* was minimized by removing the unnecessary part. A derived plasmid from pNW33N was constructed by removing the remaining part of ampicillin resistance gene amp^R. The resulting vector was named pJZ04n (Figure 4.2a), having a reduced size of 3.5 kb. It included the essential components as an *E. coli-Geobacillus* shuttle vector: chloramphenicol resistance gene $chlor^R$, multiple cloning sites, *repB* gene for replication in *Geobacillus*, and origin of replication for *E. coli*. The pJZ04n was used for protein expression and pathway construction in the following work.

To construct a synthetic pathway in a strain, strong promoters were needed to increase transcription of the genes and improve the production of the valuable biochemical. In Chapter 3, only P_{recA} was developed for protein expression. Here, different constitutive promoters were screened and characterized for pathway construction in *G. thermoglucosidasius*. In another study

of isobutanol production by Lin *et al* [7], some promoters were characterized in this strain by using thermostable *lacZ* gene as a reporter, including native glycolytic promoters and ones from *B. subtilis* and *G. thermodenitrificans*. It found P_{ldh} from *G. thermodenitrificans* was the strongest among them. Here, the P_{ldh} from *G. thermodenitrificans* together with other two active promoters in Lin's study were compared with other native promoters from *G. thermoglucosidasius* and strong ones from *B. subtilis*. A thermostable *gfpmut3** (by Dr Kang Wu) was used here for characterizing these promoters in the strain, which had relatively high activity up to 50°C. Different promoters were cloned without their own RBS into the plasmid of pJZ48 harboring the gene of *gfpmut3** with RBS of the native *recA*. As shown in Figure 4.2b, the P_{ldh} from *G. thermodenitrificans* was still the strongest among the promoters selected in this study. The P_{ldh} from *G. thermoglucosidasius* performs similar to P_{ldh} from *G. thermodenitrificans*. Therefore, P_{ldh} from *G. thermodenitrificans* was chosen for driving expression of *R*-BDO pathway in *G. thermoglucosidasius*.

4.2.2 Construction of *R*-BDO pathway

BDO is a valuable chemical, which attracts a lot of interest to engineer its production in different microorganisms. However, there are very limited reports in thermophiles [8,9,10]. Thus, it was selected as the target product in this thesis. Besides, BDO has low toxicity to microorganisms [11], which was also confirmed in *G. thermoglucosidasius* 95A1. The BDO toxicity assay for this strain was performed by detecting cell growth in TMLM with different concentration of BDO (from Sigma) ranging from 0~100 g/L. From the result (Figure 4.3), 95A1 tolerated BDO in medium up to 60 g/L, which is much higher than ethanol tolerance of this strain (12.5 g/L from Chapter 3). Little growth could be observed when the BDO concentration exceeded 80 g/L.

Based on the genomic DNA analysis, *G. thermoglucosidasius* 95A1only has the pathway to produce *R*-BDO (Figure 4.1), which has the specific applications, e.g. the use of *R*-BDO as an antifreeze agent and a low freezing point fuel. Therefore, a pathway for enhanced *R*-BDO synthesis was constructed in 95A1. In Chapter 3, it was found the main fermentation product in the wild type of this strain was lactate (Figure 3.1), so the *ldh* knockout strain (JZ01) was chosen as the start strain for *R*-BDO production (Figure 4.1). During growth of JZ01 in TMLM with 2% glucose, little *R*-BDO production was observed (Figure 4.4). To divert more carbon source to *R*-BDO synthesis, a new pathway was introduced by overexpressing heterologous acetolactate decarboxylase (*alsD*) under the P_{ldh} from *G. thermodenitrificans* (Figure 4.1). AlsD catalyzes the conversion of *R*-2-acetoin from 2-acetolactate, avoiding the requirement of oxygen. *B. licheniformis* was chosen as the donor of AlsD, because it was reported to produce *R*-BDO at high temperature (50°C) by its native pathway. However, the strain harboring the new enzyme didn't perform better on *R*-BDO production (Figure 4.4). According to the previous study of isobutanol production by Lin *et al*. [7], the native acetolactate synthase (*alsS*) had very low activity on formation of acetolactate, which was also the first key enzyme for isobutanol production. By comparison of activities, thermostabilities and specificity among different AlsS, they found the AlsS from *B. subtilis* had the highest specific activity on acetolactate formation from pyruvate without product inhibition, which was thermostable up to 60°C. This result was also consistent with other studies [12,13]. Based on the above, AlsS from *B. subtilis* was co-expressed with AlsD here to improve *R*-BDO production. As shown in Figure 4.4, this new construct significantly increased *R*-BDO production, which indicated active AlsS was required in the pathway. The expression of $alsS_{Bs}$ shifted carbon flow away from ethanol production to *R*-BDO production. The strains were cultured under aerobic or anaerobic conditions. By

comparison of production under aerobic and anaerobic conditions, aerobic culturing was preferred for *R*-BDO production but the conditions still need to be optimized, because the production decreased after its highest point.

4.2.3 Selection of thermostable enzyme for *R*-BDO production

To optimize the new pathway for *R*-BDO synthesis, more *alsD* genes were cloned and co-expressed with $alsS_{Bs}$. Three *alsD* genes were found by blasting from thermophilic bacteria, *Streptococcus thermophilus, Pelotomaculum thermopropionicum* and *Methanosaeta thermophila*. These three genes, $alsD_{St}$ (STU_RS13910), $alsD_{Pt}$ (PTH_RS12140) and $alsD_{Mt}$ (Mthe_0909) were synthesized by Integrated DNA Technologies, and replaced of $alsD_{Bl}$, respectively. By comparison of the strains harboring the plasmids with $alsS_{Bs}$ and *different alsS* (Figure 4.5), the expression of $alsD_{St}$, together with $alsS_{Bs}$, had the highest production of *R*-BDO. On the contrary, the expression of $alsD_{Pt}$ or $alsD_{Mt}$ with $alsS_{Bs}$ produced little *R*-BDO, which indicated co-expression of thermostable and active *alsS* and *alsD* resulted in improved *R*-BDO production.

Three *alsS* genes from different *Geobacillus* species (*G. caldoxylosilyticus* NBRC 107762, *G. stearothermophilus* 10 and *G. thermoglucosidasius* BGSC 95A3) were then selected and expressed separately with $alsD_{St}$ in JZ01. These strains were compared with the one carrying $alsS_{Bs}$ and $alsD_{St}$ for *R*-BDO production. As shown in Figure 4.6, the expression of *alsS* from different *Geobacillus* species produced little *R*-BDO. Therefore, *alsS* from *Geobacillus* may generally have low activity to catalyze the formation of acetolactate. The strain harboring the $alsS_{Bs}$ and $alsD_{St}$ co-expression plasmid was selected for the following study.

4.2.4 Conditions optimization for *R*-BDO production

To improve *R*-BDO production, strains were cultured in flasks and fermentation conditions were optimized, including oxygen supply, temperature, inoculation time, medium and concentration of yeast extract.

4.2.4.1 Oxygen supply

According to other studies [11,14], oxygen supply in fermentation has important effect on BDO production, because it is related to $NADH/NAD^+$ ratio in the cell. To adjust the oxygen supply during fermentation, different volumes of medium together with different agitation were tested. As shown in Figure 4.7a, the optimal production was achieved at agitation speed of 150 rpm when 60 mL or 80 mL medium was filled in 125 mL flasks. At last, 80 mL/125mL at 150 rpm was selected because of higher yield.

4.2.4.2 Temperature and inoculation time

The effect of temperature on the production of *R*-BDO was studied, because it was closely related to enzyme activity and cell growth. $AlsS_{Bs}$ was thermostable below 60°C while optimal temperature for *G. thermoglucosidasius* was 60°C. Experiments were then performed at 52°C, 55°C, and 60°C, respectively. As shown in Figure 4.7b, the production of *R*-BDO was similar at 52°C and 55°C. It decreased significantly at 60°C because enzyme activity reduced. Based on the result, 55°C was then selected for future investigations.

Different inoculation time was tested, which was correlated with growth time of seed culture. As seen in Figure 4.7c, difference in production was small when the time was range from 10 h to 15 h.

4.2.4.3 Different medium and yeast extract

In previous fermentations, minimal medium of TMLM was used for *R*-BDO production. Another minimal medium (named modified ASYE in this study) was tested and compared to TMLM, which was used for isobutanol production in *G. thermoglucosidasius* in Lin's paper [7]. As a result (Figure 4.8a), modified ASYE was more suitable for *R*-BDO production possibly because of its better buffer system. At the end of fermentation, the final pH was kept around 6.5 in modified ASYE with HEPES buffer, while the final pH dropped below 6 in TMLM with phosphate buffer (both had pH~7 before fermentation).

It was reported that increased concentration of yeast extract improved, because BDO production was considered to be related to cell growth [15]. Yeast extract is used as an organic nitrogen source to enhance cell growth, so its addition can expedite cell growth. Different concentration (from 0.1% to 1%) of yeast extract was added into medium for *R*-BDO production (Figure 4.8b). As expected, more *R*-BDO produced when more yeast extract was added, together with increased glucose consumption and ethanol formation. 7.2 g/L of *R*-BDO produced when 0.7% yeast extract was added after 48 h, at 55°C (Figure 4.8c).

As a result, the optimal fermentation conditions for *R*-BDO production was as follows: fermentation was performed in 125 mL flask with 80 mL modified ASYE (with 2% glucose and 0.7% yeast extract) inoculated by 12 h old seed culture for 48h, at 55°C, 150 rpm.

R-BDO production was also carried out in bioreactor to improve the titer. However, no promising result was obtained (Figure 4.9). Fermentation conditions are needed to be optimized furtherly.

4.2.5 Inactivation of Adhs

From the above results, *R*-BDO production was competed with ethanol formation, because both of them were converted from pyruvate. Therefore, reduced ethanol production might drive more carbon source to synthesize more *R*-BDO. To verify this hypothesis, Adhs in *G. thermoglucosidasius* 95A1 were inactivated separately. Four *adh* genes were found in the genome, AOT13_RS06455, AOT13_RS03700, AOT13_RS15005, AOT13_RS03325, which were potential functional for ethanol formation. As a result from Figure 4.10, inactivation of Adh did decrease ethanol formation but didn't help improve *R*-BDO production. Deletion of *adh* increased *R*-BDO yield but affected cell growth.

4.3 Discussion

In this chapter, a new pathway was constructed for *R*-BDO production in *G. thermoglucosidasius*. New strong promoters were tried to develop for protein expression system in this strain by screening different native promoters together with strong ones from *B. subtilis*. Same as results in another paper [7], P_{ldh} from *G. thermodenitrificansius* was the strongest, although they were from different strains. Different from their reporter, a thermostable GFPmut3* (done by Dr. Kang Wu) was used for charactering different promoters. This fluorescence protein is a mutant of GPFmut3 with three mutations (F99S, M153T and V163A) and it works in *G. thermoglucosidasius* up to 55°C. Another green fluorescence protein (sfGFP) was developed recently in *G. thermoglucosidasius*, which could work at 60°C [16].

The wild type of *G. thermoglucosidasius* 95A1 produced little of *R*-BDO. The reason should be the low activity of native AlsS for acetolactate synthesis from two pyruvate molecules, which is key enzyme for BDO production. Few active AlsS from thermophilic bacteria were reported [7]. AlsS from different *Geobacillus* also had low activity of *R*-BDO production. The only one,

thermostable above 50°C, was from *S. thermophilus*, but it had lower specific activity than AlsS from *B. subtilis* [7]. The thermostability and specific activity of $AlsS_{Bs}$ indicated protein from mesophilic bacteria may also be thermostable. Besides, its thermostablity can be improved by direct evolution strategy. A screened strain of *Geobacillus sp*. XT15 was reported to have native ability of exporting BDO at 55°C [10]. Its AlsS or the one from its neighbors could be tested for future work. The endogenous pathway for *R*-BDO production required oxygen. Heterologous expression of AlsD avoided this requirement, because it may decrease the yield of BDO. The last enzyme for *R*-BDO synthesis was *2R*, *3R*-butanediol dehydrogenase (BdhA), which was not overexpressed in this study. It was considered to be active enough for *R*-BDO under the current status, because little acetoin was accumulated during fermentation. Another reason was that more protein overexpression might affect growth rate due to the protein burden [17]. Therefore, the pathway constructed for *R*-BDO consisted of AlsS from *B. subtilis* and AlsD from *S. thermophilus* under the promoter of *ldh* from *G. thermodenitrificansius*.

Fermentation conditions were optimized for efficient *R*-BDO production. Oxygen level in the cell affected much on this production [7,18,19,20]. According to those studies, limited oxygen conditions promoted BDO synthesis, which was consistent with our results. NADH was the important cofactor of BdhA in the BDO pathway (Figure 4.1). NADH level was high under anaerobic condition, which was accumulated in glycolysis. However, more carbon source flowed to ethanol production, since this pathway consumed more NADH. Both of *R*-BDO and ethanol synthesis were using pyruvate as substrate. The heterologous pathway of *R*-BDO needed one NADH for one *R*-BDO formation, while ethanol pathway needed one pyruvate together with two NADH. This could explain the strain of $\Delta ldh/alsS_{Bs}/alsD_{Bl}$ performed worse under anaerobic

conditions (Figure 4.4). When the oxygen level was high, more carbon source would be utilized for cell growth and NADH would be too low to produce ethanol and BDO (Figure 4.7a).

Deletion of *adh* increased BDO yield because of carbon source shifting from ethanol production (Figure 4.10). However, it didn't increase the production of *R*-BDO, because it also decreased growth rate. NAD^+ and NADH are redox cofactors involving in various enzymatic reactions in the cell, so they play an essential role in cellular metabolism. Maintenance of redox balance is required fundamentally for cell growth and cellular metabolism [21,22]. The inactivation of ethanol pathway might imbalance NAD^+ and NADH levels in the cell, because extra NADH accumulated without consumption by Adh. Therefore, maintaining redox balance is an important factor to be concerned in metabolic engineering. This imbalance can be resolved by expression of NADH oxidase, NADH kinase or by modulating the cofactor specificity of pathway enzymes [22,23,24]. This can be a direction for further study of *R*-BDO production in *G. thermoglucosidasius*.

4.4 Conclusion

In conclusion, we engineered *G. thermoglucosidasius* 95A1 to efficiently produce a valuable chemical of *R*-BDO at 55°C. This is the first report of enantiopure *R*-BDO production in thermophiles at a high temperature. The production was improved through optimization of fermentation conditions. To improve the production furtherly, more thermostable and active enzymes in the pathway are needed, which can be found in nature or engineered by direct evolution. Besides, the strain can be metabolic engineered together with NAD^+ or NADH regeneration system to enhance the titer and yield. This chapter confirmed *G. thermoglucosidasius* as a good producer for different biofuels and valuable chemicals.

4.5 Material and methods

4.5.1 Plasmid construction

All plasmids used in this study are listed in Table 4.1. All primers used in this study are listed in Table 4.2. The *G. thermoglucosidasius* 95A1 shuttle plasmid pJZ04n was constructed by assembling two fragments PCR amplified from pNW33N. One fragment containing *ori* and 5' end of *repB* genes were amplified using the primers of JZ300F and JZ324R. The other containing $chlor^R$ and the rest of *repB* genes were amplified using JZ325F and JZ303R. The two DNA fragments were then assembled by the Gibson Assembly Cloning Kit (NEB).

The GFPmut3* plasmid of pJZ48 ($chlor^R$) without any promoters was constructed for promoter screening. The *gfpmut3** gene was PCR amplified from pKW1051(constructed by Kang Wu) using the primers of JZ234F2 and JZ235R2 and the *recA* terminator was PCR amplified from *G. thermoglucosidasius* 95A1 genomic DNA using the primers of JZ213F2 and JZ214R2. The two fragments were then assembled using Gibson Assembly Cloning Kit to obtain the plasmid pJZ48. Different promoters were PCR amplified using their respective primer pairs and then cloned into pJZ48 using the SacI and XbaI restriction sites.

The *alsD* expression plasmid pJZ66 ($chlor^R$) was constructed as follows. The *alsD* gene was PCR amplified from *B. licheniformis* ATCC 9945A genomic DNA using the primers of JZ432F and JZ433R. The resulting fragment was then cloned into the XhoI and AgeI sites of the plasmid pJZ48 with the promoter of *ldh* from *G. thermodenitrificans*. The *alsS* and *alsD* expression plasmid pJZ67b ($chlor^R$) was constructed in similar manner of pJZ48 using the primer pairs JZ431F2/ JZ437R to PCR amplify the *alsS* gene from *B. subtilis* 168 genomic DNA and JZ438F/ JZ439R to amplify the *alsD* gene from *B. licheniformis* ATCC 9945A genomic DNA. The gblocks gene fragments (Integrated DNA Technologies) for different *alsD* genes from *S.*

thermophiles, *P. thermopropionicum*, *M. thermophila* were synthesized and replaced of the *alsD* gene in pJZ67b using the SpeI and AgeI restriction sites.

The *G. thermoglucosidasius* 95A1 *adh1* knockout plasmid pJZ72 (spc^R) was constructed as follows. The upstream fragment (700 bp) was PCR amplified from genomic DNA using the primers JZ466F and JZ467R and the downstream fragment (700 bp) was PCR amplified from genomic DNA using the primers JZ468F and JZ469R. The two DNA fragments were then then assembled by Gibson Assembly into the EcoRI and BamHI sites of the plasmid pJZ04s to obtain pJZ72. The other three *adh* knockout plasmids of pJZ73, pJZ75 and pJZ76 were constructed in the same manner for *adh2*, *adh3* and *adh4* knockouts, respectively.

4.5.2 Fermentations

Seed culture was prepared in 125 mL flasks with 20 mL TGP. The culture was inoculated with a single colony from a fresh plate and grown aerobically at 55°C, 250 rpm for 12 hour. Fermentation cultures were carried out in 125 mL flasks containing 80 mL freshly prepared modified ASYE with 2% glucose and yeast extract. They were inoculated with the seed culture to an initial OD_{600} of 0.1 and grown at 55°C, 125 rpm for about 48 h (conditions changed as indicated). 0.5 mL of sample was taken at time intervals as indicated and analyzed by HPLC.

Bioreactor experiments were carried out using a New Brunswick BioFlo/CelliGen 115 fermenter with 0.75 L working volume. The temperature was controlled at 55°C and the pH of the culture was maintained at 7.0 by addition of 4N NaOH or 4N H_2SO_4. Seed culture was prepared as above but using 500 mL flasks filled with 100 mL TGP. Modified ASYE was used as medium for the fermentation with 1% yeast extract and 3% glucose as carbon source. Fermentation started with a 5% inoculation of seed culture. The bioreactor was operated with agitation of 150 rpm and aeration of 0.5 vvm. Antifoam 204 (Sigma) was added as required.

4.6 Tables

Table 4.1 Plasmids and strains used in this study

Plasmid or strain	Characteristics	Source or reference
Plasmids		
pNW33N	Cm^R; *E. coli-Bacillus* shuttle vector	BGSC
pKW1051	pNW33n containing *gfpmut3**	Lab stocks
pJZ04n	Cm^R; *E. coli-Bacillus* shuttle vector; pNW33N derivative	This study
pJZ04s	Spc^R; *E. coli-Bacillus* shuttle vector; pJZ04 derivative	The 3rd Chapter
pJZ48	pJZ04n containing *gfpmut3** with RBS and terminator of *recA* from *G. thermoglucosidasius* 95A1	This study
pJZ66	*alsD* from *B. licheniformis* ATCC 9945A replaces of *gfpmut3** in pJZ48 with P*ldh* from *G. thermodenitrificans*	This study
pJZ67b	*alsS* from *B. subtilis* and *alsD* from *B. licheniformis* replace of *alsD* in pJZ66	This study
pJZ67c	*alsD* from *Streptococcus thermophiles* LMG 18311 replaces of *alsD* in pJZ67b	This study
pJZ67d	*alsD* from *Pelotomaculum thermopropionicum* replaces of *alsD* in pJZ67b	This study
pJZ67e	*alsD* from *Methanosaeta thermophila* replaces of *alsD* in pJZ67b	This study
pJZ67f	P_{rpls} from *G. thermoglucosidasius* 95A1 replaces of promoter in pJZ67c	This study
pJZ67g	P_{veg} from *Bacillus subtilis* 168 replaces of promoter of pJZ67c	This study
pJZ67h	P_{lepA} from *Bacillus subtilis* 168 replaces of promoter of pJZ67c	This study
pJZ67k	P_{ligG} from *Bacillus subtilis* 168 replaces of promoter of pJZ67c	This study
pJZ72	pJZ04s containing truncated *adh1* (AOT13_RS06455)	This study
pJZ73	pJZ04s containing truncated *adh2* (AOT13_RS03700)	This study
pJZ75	pJZ04s containing truncated *adh3* (AOT13_RS15005)	This study
pJZ76	pJZ04s containing truncated *adh4* (AOT13_RS03325)	This study
Strains		
G. thermoglucosidasius BGSC 95A1	Wild type, $DSM2452^T$	BGSC
Geobacillus thermodenitrificans 94A1	Source of P_{ldh}	BGSC
Bacillus subtilis 168	Source of *alsS*, *alsD*, P_{veg}, P_{lipA} and P_{ligG}	BGSC

Table 4.1 (cont.)

Plasmid or strain	Characteristics	Source or reference
B. licheniformis 9945A	Source of *alsD*	BGSC
G. caldoxylosilyticus NBRC 107762	Source of *alsS*	BGSC
G. stearothermophilus 10	Source of *alsS*	BGSC
G. thermoglucosidasius BGSC 95A3	Source of *alsS*	BGSC
JZ01	*Δldh* variant of *G. thermoglucosidasius* 95A1	The 3rd chapter
JZ09	*ΔldhΔadh1* variant of *G. thermoglucosidasius* 95A1	This study
JZ10	*ΔldhΔadh2* variant of *G. thermoglucosidasius* 95A1	This study
JZ11	*ΔldhΔadh3* variant of *G. thermoglucosidasius* 95A1	This study
JZ12	*ΔldhΔadh4* variant of *G. thermoglucosidasius* 95A1	This study
JZ13	*Δldh Δadh5* variant of *G. thermoglucosidasius* 95A1	This study

Table 4.2 Primers used in this study

Primer	Sequence	Plasmid
JZ300F	TAGCGATTTTCTTTTCTCTCCATGGACGCACTTTTCCACTTTTTGTCTTGTCCAC	pJZ04n
JZ303R	CCATGGAGAGAAAAGAAAATCGCTAATGTTG	
JZ324R	ATCCCCGGGTACCGAGCTCGAATTCTAACTGTCAGACCAAGTTTACTCATATATAC	
JZ325F	GAATTCGAGCTCGGTACCCGGGGATCCTC	
JZ213F2	AAACCGGTGCGAGAGAGGGACTTTGTTTTTC	pJZ48
JZ214R2	TATGACCATGATTACGCCAAGCTTGCATGCACGCGCCCCCATTTTGCC	
JZ234F2	TTCGAGCTCGGTACCCGGGGATCCTCTAGAAAGGAGGAGTCTCGAGATGCGTAAAGGAGAAGAAC	
JZ235R2	ACAAAGTCCCTCTCTCGCACCGGTTTATTATTTGTATAGTTCATC	
JZ432F	GCCTCGAGATGAAAAGTGCAAGCAAAC	pJZ66
JZ433R	TAACCGGTTTATTACTCGGGATTGCCTTCGG	
JZ431F2	GAAAGGAGGAGTCTCGAGTTGACAAAAGCAACAAAAGAAC	pJZ67b
JZ437R	TTACTAGTTTACTAGAGAGCTTTCGTTTTCATG	
JZ438F	GCTCTCTAGTAAACTAGTAAGGAGGAGTTTGTTAATGAAAAGTGCAAGCAAAC	
JZ439R	AACAAAGTCCCTCTCTCGCACCGGTAAGACGTCTTATTACTCGGGATTGCCTTCGG	
JZ466F	AAACTTGGTCTGACAGTTAGAATTCGCAATGTCTCCGCGAACTTG	pJZ72
JZ467R	CCGGTTGTCAACGTAATGTTGCGTATAACCCTTTCTTGCAATTTTC	
JZ468F	TACGCAACATTACGTTGACAACCGGTCTTG	
JZ469R	CCTGCAGGTCGACTCTAGAGGATCCAGTCCATTGCCCCACTCGTTTTG	
JZ470F	AAACTTGGTCTGACAGTTAGAATTCTCCAAATCCGACTGCTCTTTC	pJZ73
JZ471R	GCATGCGCTCCGCCCACGTTTTGCTGCTCAAACAATATTCACATTC	
JZ472F	AGCAAAACGTGGGCGGAGCGCATGCCGCC	
JZ473R	CCTGCAGGTCGACTCTAGAGGATCCGAAAAAATAAAAACGGAAAC	
JZ496F	AAACTTGGTCTGACAGTTAGAATTCTTGCCGTGAAAATTTTCGATG	pJZ75
JZ497R	AAAGTATCTTCTGCGAGCCTCATGCAACCCCCTCCTTCAATATTTC	
JZ498F	AGGCTCGCAGAAGATACTTTAAAAGAGATTGCG	
JZ499R	CCTGCAGGTCGACTCTAGAGGATCCTCTTGCAACAATCTTGCTAATTC	
JZ504F	AAACTTGGTCTGACAGTTAGAATTCTGCCACATTTTACGGCCAATCCGTG	pJZ76

Table 4.2 (cont.)

Primer	Sequence	Plasmid
JZ505R	TTAAACGCTTGGCGATAAATCTCATCAATCATTTTTGCTAC	
JZ506F	ATTTATCGCCAAGCGTTTAAAGGAG	
JZ507R	CCTGCAGGTCGACTCTAGAGGATCCTTATCAACGCTTCCAATGGC	

4.7 Figures

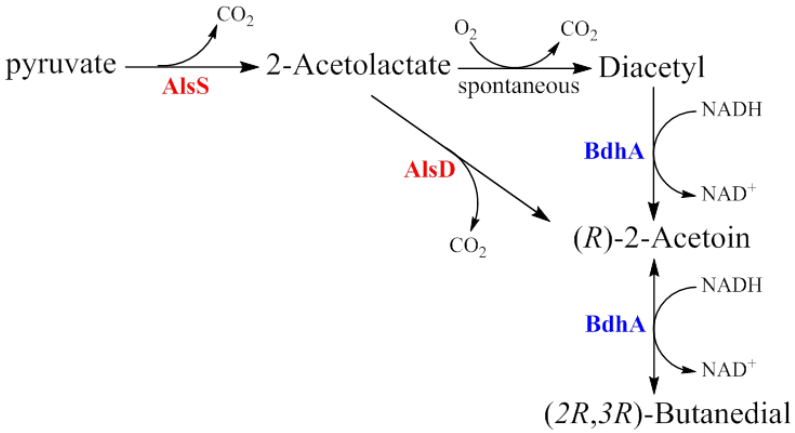

Figure 4.1 Metabolic pathway for the synthesis of (*2R,3R*)-butanediol in the engineered *G. thermoglucosidasius* 95A1. Genes in red represent the heterologous enzymes for *R*-BDO synthesis, while the blue one indicates the endogenous gene (native AlsS also exists in this strain). AlsS, acetolactate synthase; AlsD, acetolactate decarboxylase; BdhA, *2R, 3R*-butanediol dehydrogenase.

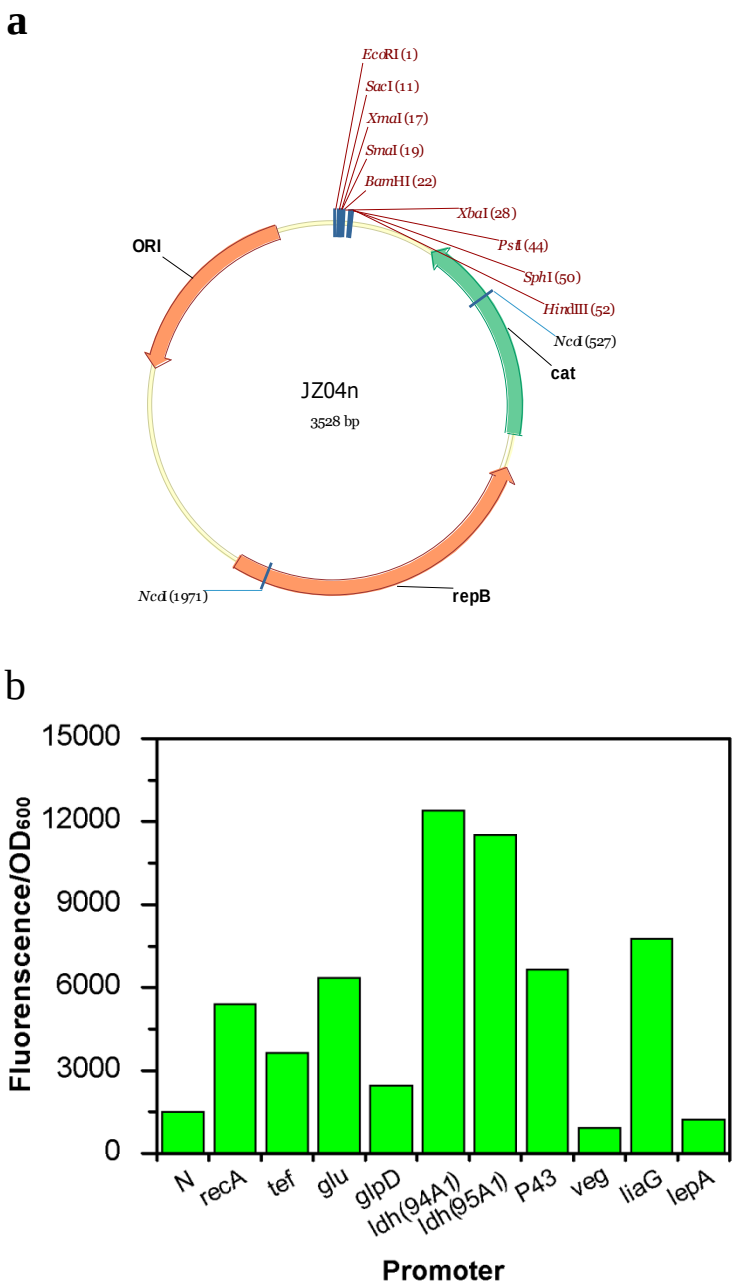

Figure 4.2 Shuttle vector construction and promoters characterization. a, map of pJZ04n, a derived vector from pNW33N without the remaining part of *amp*^{R} gene. b, different promoters were characterized by fusing with gfpmut3*. Strains harboring these plasmids were cultured at 50°C. N, negative control, wild type with empty plasmid; *recA*, recombinase (AOT13_RS09475); *tef*, transcription elongation factor (AOT13_RS17210); *glu*, PTS mannose transporter subunit II (AOT13_RS11640); *glpD*, glycerol-3-phosphate dehydrogenase (AOT13_RS15360); ldh (94A1), lactate dehydrogenase from *G. thermodenitrificans* 94A1; ldh (95A1), lactate dehydrogenase; P43, veg, liaG and lepA were from *B. subtilis* 168.

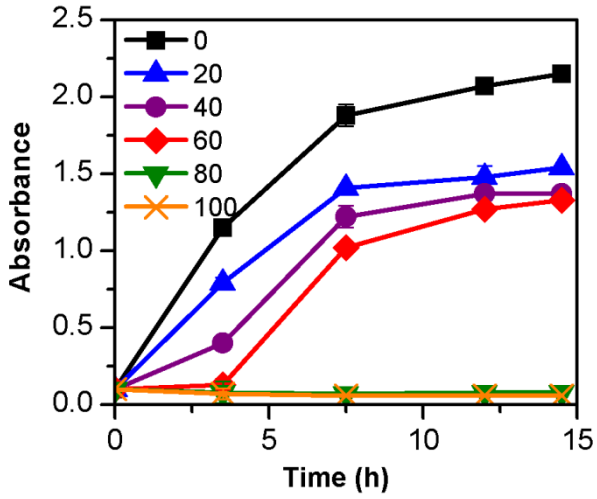

Figure 4.3 BDO tolerance of *G. thermoglucosidasius* 95A1. Cells were cultured in TMLM with 1% glucose and different concentrations of BDO from 0 to 100 g/L. The figures show the means and standard deviations obtained from two experiments.

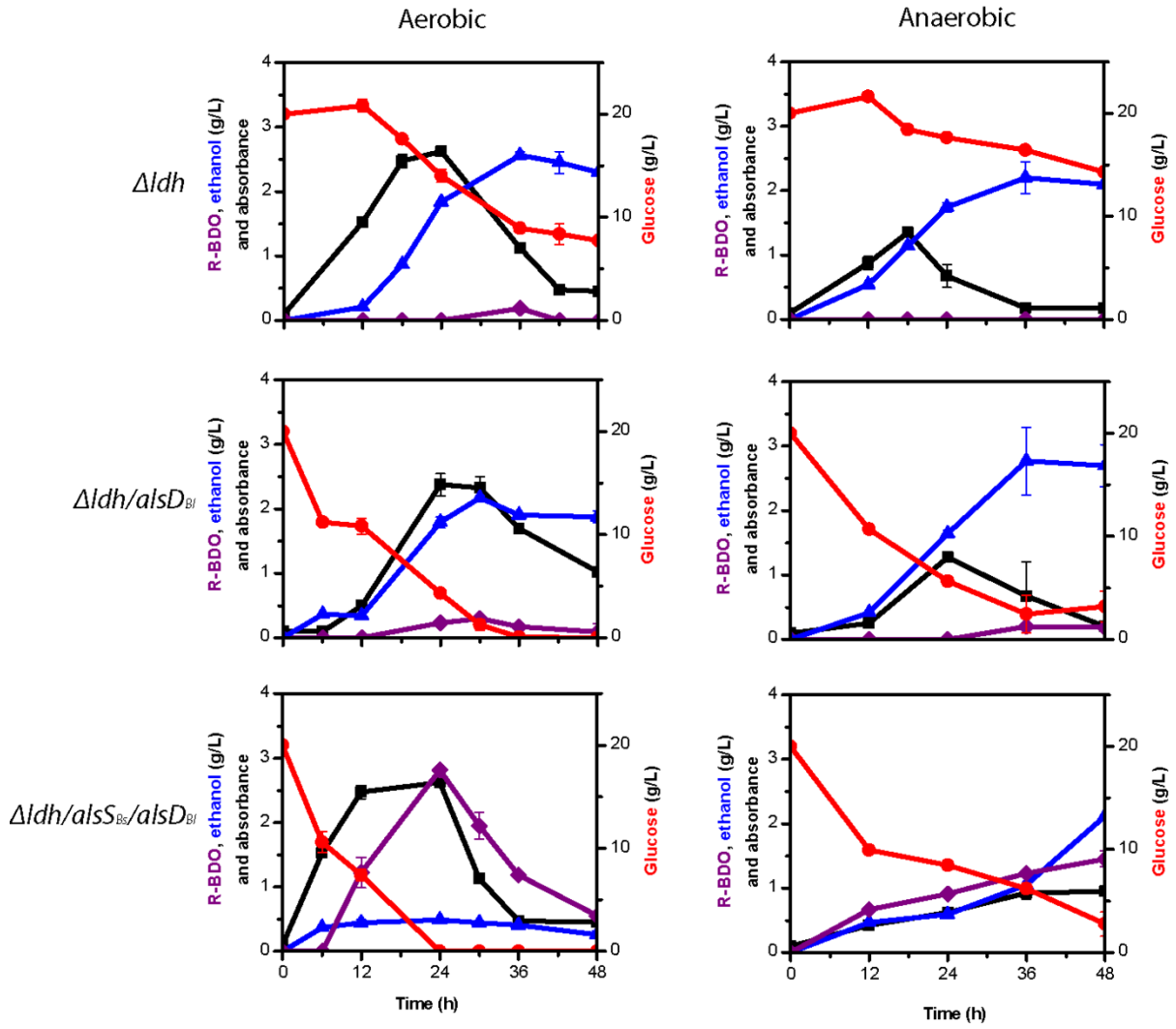

Figure 4.4 *R*-BDO production in engineered *G. thermoglucosidasius* 95A1 with different combination of enzyme. Cells were cultured in TMLM with 2% glucose, 52°C, 150 rpm. Aerobic culture was carried out in glass tubes and anaerobic culture was carried out in sealed falcon tubes as same as that for ethanol production (Chapter 3). *alsS*$_{Bs}$, *alsS* from *B. subtilis*; *alsD*$_{Bl}$, *alsD* from *B. licheniformis*. The figure shows the means and standard deviations obtained from three experiments.

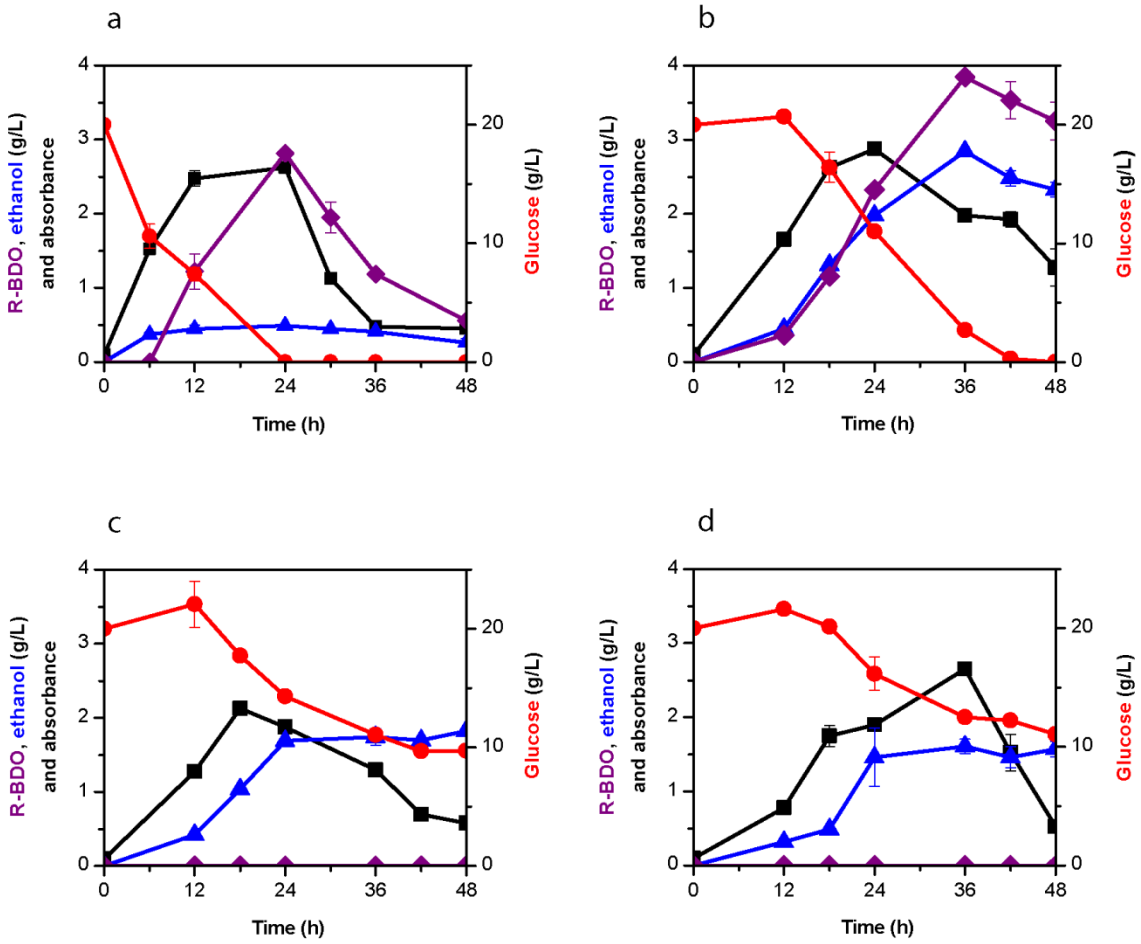

Figure 4.5 *R*-BDO production in engineered *G. thermoglucosidasius* 95A1 with different *alsD*. Cells were cultured aerobically in glass tubes with TMLM with 2% glucose and 0.1% yeast extract, 52°C, 150 rpm. a, *Δldh*/*alsS*$_{Bs}$/*alsD*$_{Bl}$; b, *Δldh*/*alsS*$_{Bs}$/*alsD*$_{st}$; c, *Δldh*/*alsS*$_{Bs}$/*alsD*$_{Pt}$; d, *Δldh*/*alsS*$_{Bs}$/*alsD*$_{Mt}$. *alS*$_{Bs}$, *alsS* from *B. subtilis*; *alsD*$_{Bl}$, *alsD* from *B. licheniformis*; *alsD*$_{St}$, *alsD* from *Streptococcus thermophilus*; *alsD*$_{Pt}$, *alsD* from *Pelotomaculum thermopropionicum*; *alsD*$_{Mt}$, *alsD* from *Methanosaeta thermophila*. The figure shows the means and standard deviations obtained from three experiments.

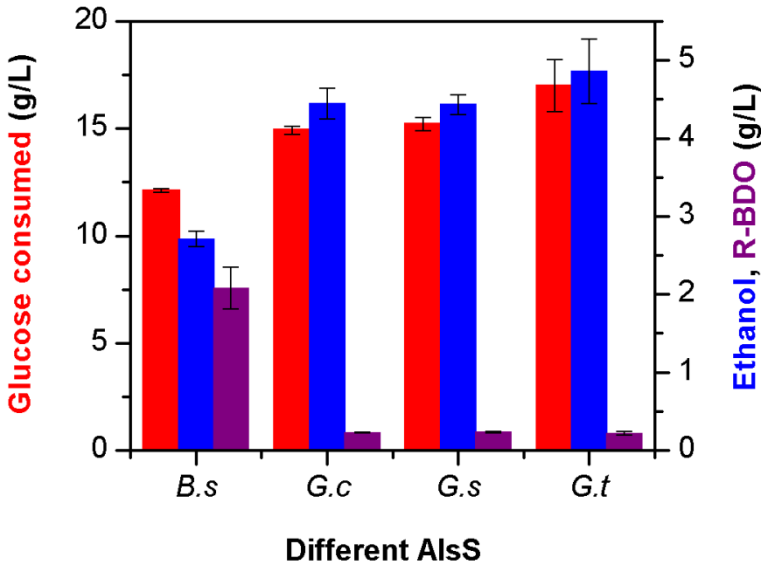

Figure 4.6 *R*-BDO production in engineered *G. thermoglucosidasius* 95A1 with different *alsS*. Cells were cultured aerobically in glass tubes with TMLM with 2% glucose and 0.1% yeast extract, 52°C, 150 rpm for 24 h. $alsD_{St}$ was co-expressed with *alsD* from different *Geobacillus* species. *B.s*, *B. subtilis*; *G.c*, *G. caldoxylosilyticus*; *G.s*, *G. stearothermophilus*; *G.t*, *G. thermoglucosidasius*. The figure shows the means and standard deviations obtained from three experiments.

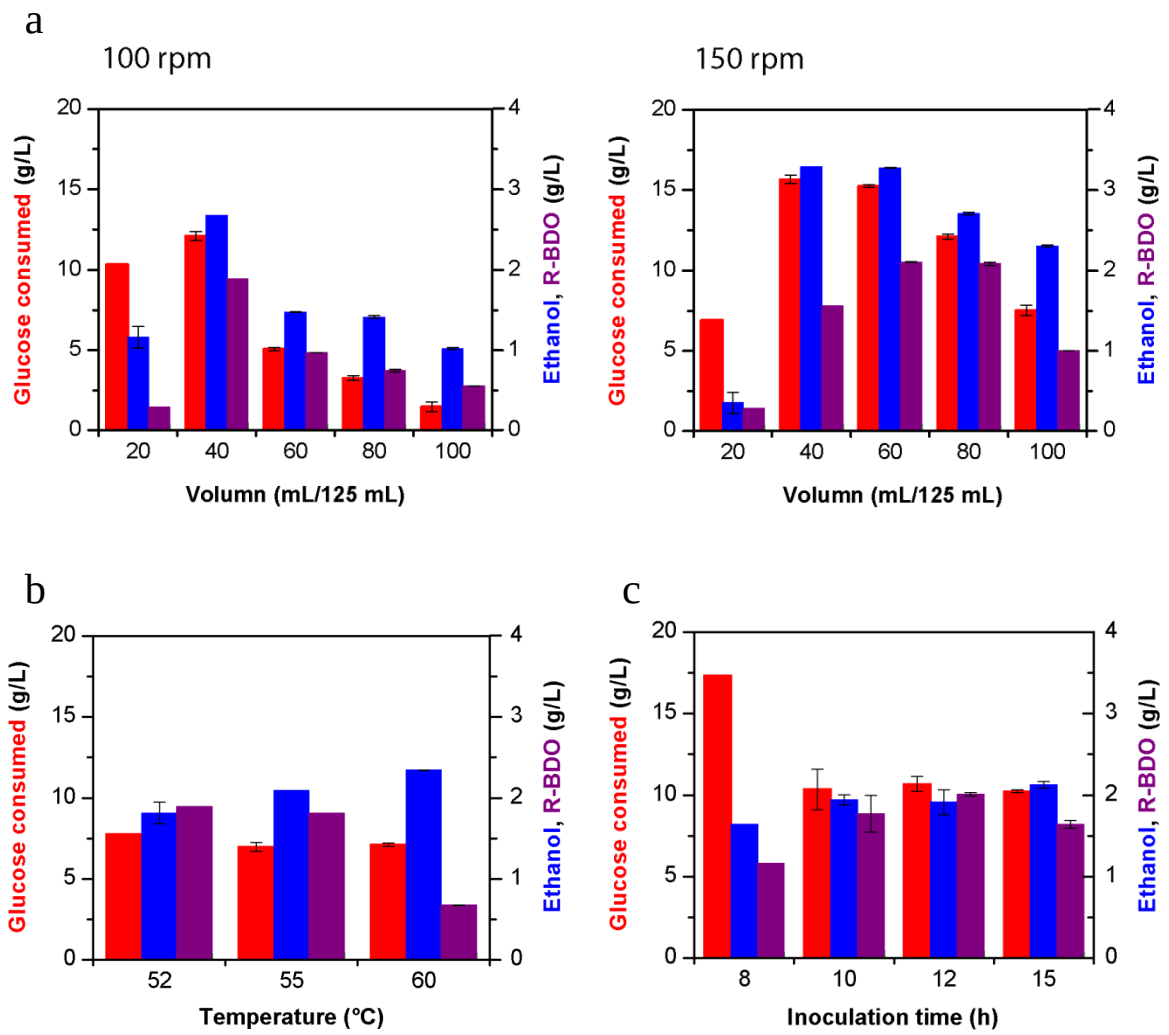

Figure 4.7 Fermentation conditions optimization for *R*-BDO production. Cells were cultured aerobically in 125 mL flask filled with TMLM with 0.1% yeast extract with 2% glucose for 24 h. a, optimization of agitation and medium volume. Cells were cultured at 52°C. b, temperature optimization. Cells were cultured aerobically in 125 mL flask filled with 80 mL medium at agitation speed of 150 rpm. c, different inoculation time (=fermentation time of seed culture). Cells were cultured aerobically in 125 mL flask filled with 80 mL medium at 52°C, 150 rpm. The figure shows the means and standard deviations obtained from three experiments.

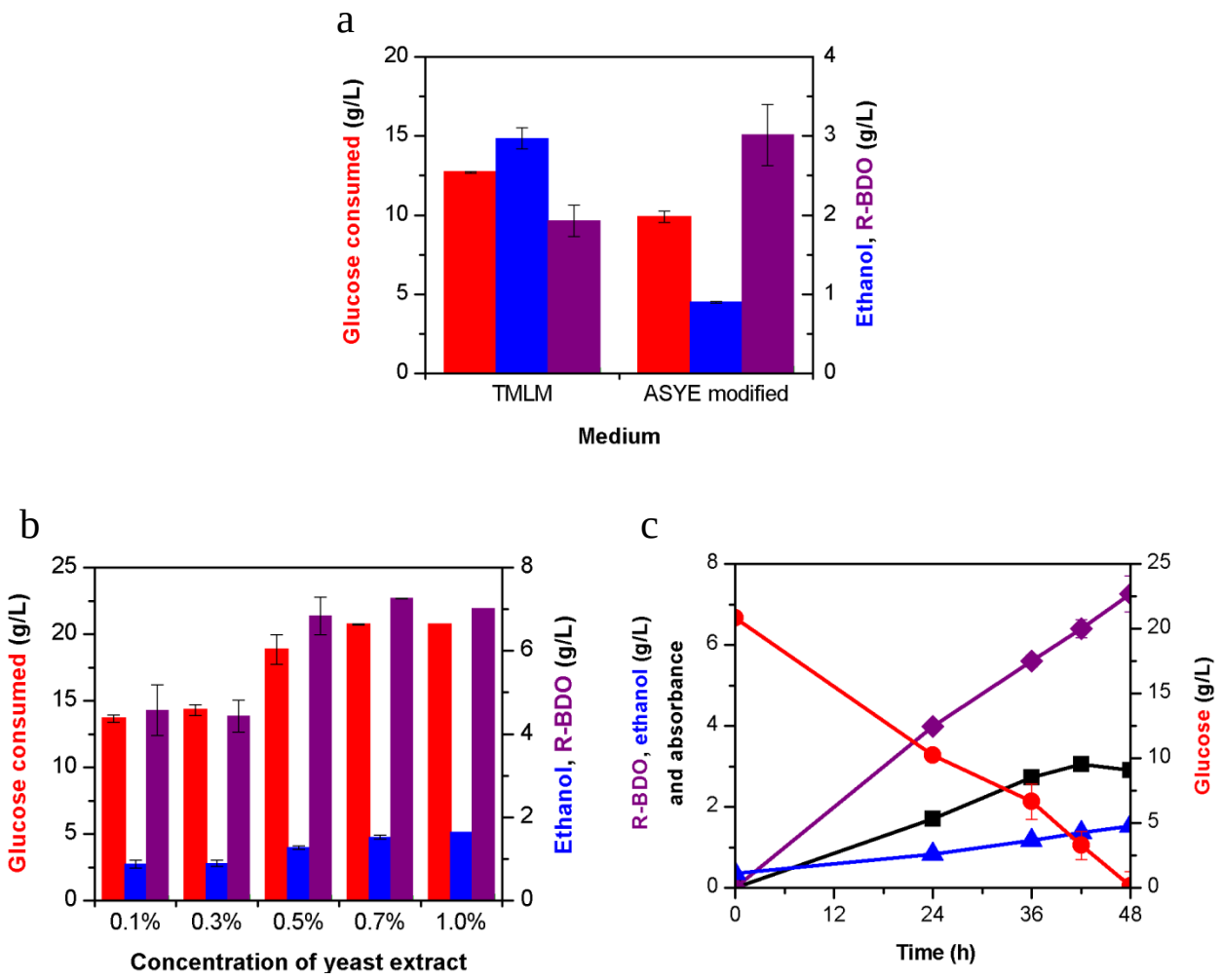

Figure 4.8 Fermentation conditions optimization for *R*-BDO production. Cells were cultured aerobically in 125 mL flask filled with 80 mL minimal medium with 2% glucose at 55°C, 150 rpm. a, different medium. Cells were cultured in medium with 0.1% yeast extract for 24 h; b, effect of different concentration of yeast extract on *R*-BDO production. Cells were cultured aerobically in modified ASYE with 2% glucose for 48 h. c, time course of *R*-BDO production in medium with 0.7% yeast extract. The figure shows the means and standard deviations obtained from three experiments.

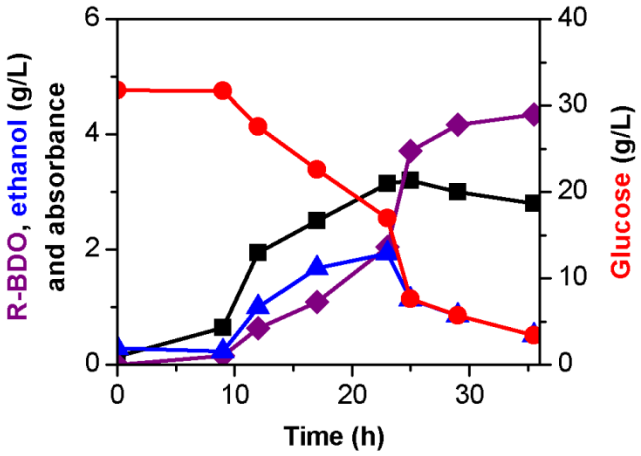

Figure 4.9 Batch growth of engineered *G. thermoglucosidasius* 95A1 in bioreactors for *R*-BDO production. Modified ASYE was used with 3% glucose and 0.7% yeast extract.

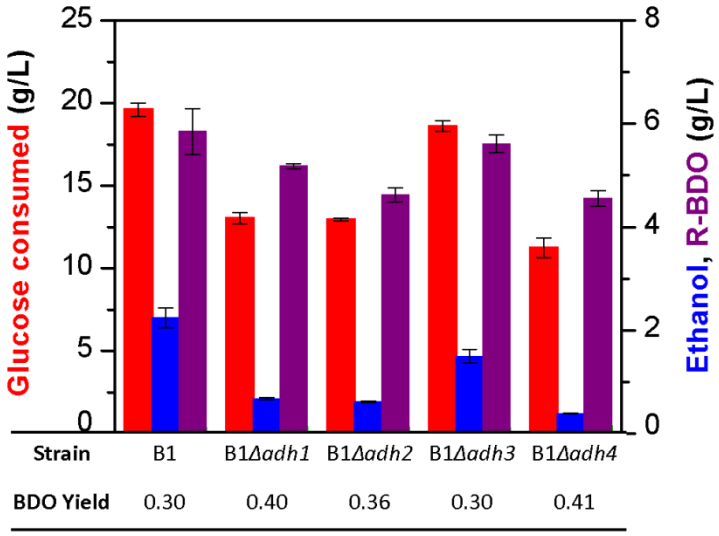

Figure 4.10 Effect of *adh* deletion on *R*-BDO production. Cells were cultured aerobically in 125 mL flask filled with 80 mL modified ASYE with 2% glucose and 0.1% yeast extract for 48 h. The figure shows the means and standard deviations obtained from three experiments.

4.8 References

1. Celińska E, Grajek W (2009) Biotechnological production of 2,3-butanediol—Current state and prospects. Biotechnology advances 27: 715-725.
2. Ji X-J, Huang H, Ouyang P-K (2011) Microbial 2,3-butanediol production: A state-of-the-art review. Biotechnology advances 29: 351-364.
3. Ji XJ, Liu LG, Shen MQ, Nie ZK, Tong YJ, et al. (2015) Constructing a synthetic metabolic pathway in Escherichia coli to produce the enantiomerically pure (R, R) ‐ 2, 3 ‐ butanediol. Biotechnology and Bioengineering 112: 1056-1059.
4. Yang T, Rao Z, Hu G, Zhang X, Liu M, et al. (2015) Metabolic engineering of Bacillus subtilis for redistributing the carbon flux to 2,3-butanediol by manipulating NADH levels. Biotechnology for Biofuels 8.
5. Lian J, Chao R, Zhao H (2014) Metabolic engineering of a Saccharomyces cerevisiae strain capable of simultaneously utilizing glucose and galactose to produce enantiopure (2R,3R)-butanediol. Metabolic Engineering 23: 92-99.
6. Kim S-J, Seo S-O, Jin Y-S, Seo J-H (2013) Production of 2,3-butanediol by engineered Saccharomyces cerevisiae. Bioresource technology 146: 274-281.
7. Lin PP, Rabe KS, Takasumi JL, Kadisch M, Arnold FH, et al. (2014) Isobutanol production at elevated temperatures in thermophilic Geobacillus thermoglucosidasius. Metabolic Engineering 24: 1-8.
8. Li L, Li K, Wang K, Chen C, Gao C, et al. (2014) Efficient production of 2,3-butanediol from corn stover hydrolysate by using a thermophilic Bacillus licheniformis strain. Bioresource technology 170: 256-261.
9. Wang Q, Chen T, Zhao X, Chamu J (2012) Metabolic engineering of thermophilic Bacillus licheniformis for chiral pure D ‐ 2, 3 ‐ butanediol production. Biotechnology and Bioengineering 109: 1610-1621.
10. Xiao Z, Wang X, Huang Y, Huo F, Zhu X, et al. (2012) Thermophilic fermentation of acetoin and 2, 3-butanediol by a novel Geobacillus strain. Biotechnology for Biofuels 5: 88.
11. Yang S, Mohagheghi A, Franden MA, Chou Y-C, Chen X, et al. (2016) Metabolic engineering of Zymomonas mobilis for 2, 3-butanediol production from lignocellulosic biomass sugars. Biotechnology for Biofuels 9: 189.
12. Gollop N, Damri B, Chipman D, Barak Z (1990) Physiological implications of the substrate specificities of acetohydroxy acid synthases from varied organisms. Journal of Bacteriology 172: 3444-3449.
13. Atsumi S, Hanai T, Liao JC (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. nature 451: 86-89.
14. Kim T, Cho S, Lee S-M, Woo HM, Lee J, et al. (2016) High production of 2, 3-butanediol (2, 3-BD) by Raoultella ornithinolytica B6 via optimizing fermentation conditions and overexpressing 2, 3-BD synthesis genes. PloS one 11: e0165076.
15. Wang X-X, Hu H-Y, Liu D-H, Song Y-Q (2016) The implementation of high fermentative 2, 3-butanediol production from xylose by simultaneous additions of yeast extract, Na 2 EDTA, and acetic acid. New biotechnology 33: 16-22.
16. Fong SS, Pogrebnyakov I, Jendresen CB, Nielsen AT (2017) Genetic toolbox for controlled expression of functional proteins in Geobacillus spp. PloS one 12: e0171313.
17. Snoep JL, Yomano LP, Westerhoff HV, Ingram LO (1995) Protein burden in Zymomonas mobilis: negative flux and growth control due to overproduction of glycolytic enzymes. Microbiology 141: 2329-2337.
18. Yang T, Rao Z, Hu G, Zhang X, Liu M, et al. (2015) Metabolic engineering of Bacillus subtilis for redistributing the carbon flux to 2, 3-butanediol by manipulating NADH levels. Biotechnology for Biofuels 8: 129.
19. Fu J, Wang Z, Chen T, Liu W, Shi T, et al. (2014) NADH plays the vital role for chiral pure D ‐ (−) ‐ 2, 3 ‐ butanediol production in Bacillus subtilis under limited oxygen conditions. Biotechnology and Bioengineering 111: 2126-2131.
20. Chan S, Jantama SS, Kanchanatawee S, Jantama K (2016) Process Optimization on Micro-Aeration Supply for High Production Yield of 2, 3-Butanediol from Maltodextrin by Metabolically-Engineered Klebsiella oxytoca. PloS one 11: e0161503.
21. Chen X, Li S, Liu L (2014) Engineering redox balance through cofactor systems. Trends in Biotechnology 32: 337-343.
22. Kim S, Hahn J-S (2015) Efficient production of 2, 3-butanediol in Saccharomyces cerevisiae by eliminating ethanol and glycerol production and redox rebalancing. Metabolic Engineering 31: 94-101.

23. Bao T, Zhang X, Rao Z, Zhao X, Zhang R, et al. (2014) Efficient whole-cell biocatalyst for acetoin production with NAD+ regeneration system through homologous co-expression of 2, 3-butanediol dehydrogenase and NADH oxidase in engineered Bacillus subtilis. PloS one 9: e102951.
24. Celińska E, Grajek W (2009) Biotechnological production of 2, 3-butanediol—current state and prospects. Biotechnology advances 27: 715-725.

Chapter 5 Screening Signal Peptides for Heterologous Protein Secretion in *G. thermoglucosidasius*

5.1 Introduction

Secretion is the process of exporting proteins to extracellular medium [1], which makes protein separation much easier. *Bacillus* is widely used in industry for extra-cellular proteins production because they are capable of secreting proteins at high concentration. As a close relative, thermophilic *Geobacillus* also possess this ability and are drawn attention for secretion of industrial thermostable proteins, such as amylase [2,3], lipase [4,5] and protease [6,7]. Thermostable enzymes are attractive in industry because they are generally more stable and more tolerant to temperatures and pH. By analysis of genomes and comparison with *B. subtilis*, it is found members of *Geobacillus* have the main components of Sec and Tat pathways for exporting proteins [8,9].

In our lab, *G. thermoglucosidasius* was identified capable of secreting proteins at high titers. Dr. Angel Rivera (an alumnus in our lab) examined the native capacity *G. thermoglucosidasius* to secrete protein on different substrates and found it was able to secrete grams per liter of proteins, especially on C5 sugars. He identified secreted proteins in *G. thermoglucosidasius* by shotgun proteomics, which was also reported by Holland [9]. In *G. thermoglucosidasius*, the Sec pathway is predominant in the strain which exports unfolded proteins [8]. Signal peptide (SP) is required for identification of secretory protein before exporting, which is around 30 amino acids at N-terminal of preprotein and cleavable during secretion [10]. Secretory endoglucanase *Cel5A* from *Thermotoga maritima* was expressed in *G. thermoglucosidasius* NCIMB 11955, which was secreted by fusing the encoding gene with signal peptide sequence of endo- β 1-4 xylanase from

G. thermoglucosidasius C56-YS93 [8]. The xylanase from *G. thermoglucosidasius* C56-YS93 with its native SP was overexpressed and secreted in *G. thermoglucosidasius* TM242 [9].

Since SP plays a crucial role in the secretion of heterologous protein [11,12], efficient SPs need to be screened for desired target protein separately. According to the studies of screening signal peptides for protein secretion in *B. subtilis*, no one-size-fits-all signal peptide exists for the optimal secretion of various proteins [8,13]. Therefore, signal peptides will have to be screened for different secretory protein at high levels. Here, the top 25 signal peptides from *G. thermoglucosidasius* C56-YS93 were used, which were predicted by the online SignalP 4.1 server, for heterologous secretion of thermostable proteins in *G. thermoglucosidasius* 95A1. Thermostable proteins were selected, including an α-amylase and two cellulases, as secretory targets, which have paid much attention for their important industrial purpose, such as hydrolysis in biofuel production. The thermostable α-amylase (EC 3.2.1.1) was from *G. stearothermophilus*. It hydrolyzed α-1,4-glucosidic linkages of the starch and produces reducing sugar. Therefore, it was easily to be detected by iodine test and dinitrosalicylic acid (DNS) method. The cellulase from *B. subtilis* was a thermostable endoglucanase, while the other was from thermophilic *C. bescii* and was bifunctional with both of endo- and exoglucanase activity.

5.2 Results

5.2.1 Construction of protein secretion plasmids with different SPs

SignalP is a popular signal peptide prediction program accessible online. It can predict the likelihood of a specific amino acid sequence being a signal peptide or not by calculating a discrimination score[13]. In our lab, SignalP 4.1 was used for predicting SPs in *G. thermoglucosidasius* C56-YS93, since its genome sequence was available online when this project started (done by Dr Angel Rivera). According to the scores, the top 25 SPs (listed in

Table 5.1) were picked up for screening. 25 DNA fragments encoding SPs were separately cloned into pGEM-T easy Vector for sequencing before cloning into the receiving plasmid by Golden Gate assembly (Figure 5.1a). The old cloning strategy was cloning the PCR products directly into the receiving plasmid, but it was found mutations happened frequently at the ends of SP sequence. The new strategy avoided this problem and had high cloning efficiency.

Besides, the constructs of different SPs on T vectors can be used for fusing with different secretory proteins with designed junctions (Figure 5.1b). The design in this study was introducing an alanine at N-terminal of the secretory protein if the first amino acid is not Ala. The junction at C-terminal with SP sequence was "GGCA": the first "G" was designed as the last nucleotide of SP, since it wouldn't change the amino acid; the "GCA" was for the Ala at N-terminal of protein. The junction at N-terminal was "CGAC", which are the last four nucleotides of XhoI site introduced in RBS of *recA* from *G. thermoglucosidasius* right before encoding region. The signal peptides with this design were fused with different cellulase in this chapter. The junctions between signal peptide sequence and the amyE gene were the last four nucleotides in its native RBS from *G. stearothermophilus* 9A21 and the first four nucleotides in the gene without its native SP.

5.2.2 Screening of SPs for heterologous amylase secretion

α-Amylase was firstly chosen as the target for heterologous protein secretion in *G. thermoglucosidasius* due to its important biotechnological application and convenient methods of activity detection. The truncated *amyE* gene was cloned from *G. stearothermophilus* 9A21, which encoded a thermostable α-amylase (59 kDa) with an optimal temperature of 70°C (Figure 5.2a). It was fused with DNA fragments of different SPs and expressed under its native promoter

with the RBS. The 25 resulting plasmids were all transformed successfully into *G. thermoglucosidasius* 95A1, which doesn't have amylase genes.

The efficiency of different SPs in amylase secretion was then determined by two methods. One was iodine test by spotting culture on starch plates. Secreted amylase degraded the starch around the colonies, which formed clear zones on plates treated with iodine solution. The size of clear zones indicated the secretion efficiency of SPs. The other was measuring reduced sugars in enzymatic reaction by DNS method. The efficiency of 25 SPs in amylase was screened by these two methods (Figure 5.2b,c). The results from two methods were almost consistent.

The secretion efficiency of 25 SPs showed no correlation between secretion efficiency and discrimination scores (D-scores). Higher scores correspond to higher probability of being a SP: D-scores above 0.7 usually indicate that the corresponding sequence does indeed represent a signal peptide. In contrast, D-scores between 0.5 and 0.7 show lower possibility to be a functional SP, while the sequences with values below 0.5 are usually classified as not being functional signal peptides [13]. D-scores of the 25 SPs in this study were all above 0.7 or nearly 0.7 (Table 5.1), indicating that all of them are functional. From the result of DNS method, it showed secretion efficiency was not correlated to the D-score.

The native SP of amylase from *G. stearothermophilus* 9A21 was tested as the positive control, which was set as a reference. Most SPs were found to have higher secretion efficiency than the SP of AmyE. There were 9 SPs have increased activity by more than 6-fold compared to the native SP: SP3, SP7, SP8, SP9, SP11, SP13, SP14, SP21 and SP25 (Figure 5.2c). The screening of the 25 SPs revealed strong differences in amylase activity of the culture supernatants. The SP21, SP of a putative uncharacterized protein (from GEOTH_RS18390), had

the highest secretion efficiency, which increased amylase activity by about 18-fold. By contrast, SP15 with amylase showed no secretion at all (Figure 5.2c).

A strategy of screening signal peptides for optimizing heterologous protein secretion in *G. thermoglucosidasius* was established. It was used for different protein targets, since other studies about SPs concluded no one-size-fits-all signal peptide existed for the optimal secretion of various proteins.

5.2.3 Screening of SPs for endoglucanase EglS secretion

Cellulase is another industrially important enzyme, which plays a key role in cellulosic biofuel production. As introduced in Chapter 1, cellulase divides into 3 types: endoglucanase, exoglucanase and β-glucosidase. For efficient cellulose hydrolysis, different types of cellulase are required, which act on cellulose synergistically. Since *G. thermoglucosidasius* 95A1 has β-glucosidase, it can utilize cellulose by secreting endoglucanase and exoglucanase. Although genome sequencing predicts the presence of an endoglucanase (AOT13_RS18045) in *G. thermoglucosidasius* 95A1 [8], it is not secretory and has no hydrolysis shown on CMC plates (Figure 5.3c). Generally, cellulase is desired in the form of secreted protein for degradation of cellulosic biomass in practical use.

B. subtilis was reported to have an endogenous endoglucanase, which is thermostable above 50°C [14]. The truncated gene encoding for endoglucanase EglS (60 kDa, BSU18130) from *B. subtilis* 168 was cloned with the promoter of *ldh* from *G. thermodenitrificansius* (from Chapter 4). The DNA fragments of 25 SPs were inserted into the receiving plasmid of *eglS*, which formed 25 EglS secretion plasmids with different SPs. *G. thermoglucosidasius* 95A1was transformed with the 25 EglS secretion plasmids separately and tested for extracellular endoglucanase by CMC (carboxymethyl cellulose) plates and DNS method. CMC plates are

often used for endoglucanase activity assay. If a strain is cellulolytic, it will hydrolyze CMC in the surrounding, which form a clear zone around the colony after the plate is strained with Congo red and destained with NaCl solution. The size of clear zones indicates the secretion efficiency of cellulase.

The optimal temperature of EglS was 60°C (Figure 5.3a), which was set for the enzyme reaction. The result of the activity assay by DNS method was consistent with that from CMC plates (Figure 5.3b, c). The secretion efficiency of 25 SPs for EglS has no relation with D-scores. The EglS with its native SP from *B. subtilis* was also expressed in *G. thermoglucosidasius* as the positive control. The result form CMC plates showed the native SP also had good secretion efficiency for EglS in *G. thermoglucosidasius*. The difference of screening efficiency of the 25 SPs in endoglucanase activity was not as big as that in amylase activity. 14 SPs had higher secretion efficiency than the native SP (Figure 5.3c). Interestingly, SP21 still had the highest secretion efficiency, increasing activity by less than 2-fold compared to the positive control. SP15, SP20 and SP24 showed no obvious secretion in both results of two assays (Figure 5.3c).

5.2.4 Screening of SPs for cellulase CelA secretion

Another cellulase of CelA (ATHE_RS09370) from *C. Bescii* was chosen for heterologously secretion in *G. thermoglucosidasius* 95A1. *C. Bescii* is a thermophilic bacterium that can degrade and utilize cellulose and hemicellulose. It has multiple cellulolytic enzymes, while CelA is reported to be the most highly secreted. The *C. bescii* CelA is a multi-modular enzyme, consisting of an N-terminal GH9 module, a C-terminal GH48 module, and three carbohydrate-binding modules between them (Figure 5.4). It is bifunctional: GH9 module possesses endoglucanase activity, while GH48 module possesses exoglucanase activity [15]. It has been well characterized, having an optimal temperature of 75°C [15,16,17]. It is a large polypeptide

(~190 kDa), so the GH9 and GH48 module were also expressed with three CBM. Thus, three constructs were built for heterologous secretion of CelA, GH9/CelA (115 kDa) and GH48/CelA (133 kDa), respectively (Figure 5.4). 25 SPs were fused with constructs separately to test their secretion efficiency for CelA and the two mutants.

Secretion efficiency for CelA and GH9/CelA was tested by CMC plates, the same way for EglS activity. Enzyme activity was determined by measuring the reducing sugar using pHBAH method (described in 2.3.2.2) instead of DNS method. DNS method is not sensitive enough here, which has a measuring range of 0~2 g/L glucose. The pHBAH method has a measuring range of 0~0.5 g/L glucose [18]. Enzyme activity of GH48/CelA was determined by measuring the concentration of cellobiose in enzyme reaction.

The screening result of CelA from the CMC plate was not completely consistent with that from enzyme assay (Figure 5.5). It is possibly because the substrate used for enzyme assay was PASC but not CMC. SP might influence substrate specificity, which was also reported in another study [19]. From result of CMC plate (Figure 5.5a), the average activity of CelA was lower than that of GH9/CelA, most probably because of its large size. From the enzyme assay (Figure 5.5b), SP4, SP of Peptidase P60 (from GEOTH_RS17105), had the highest secretion efficiency for CelA secretion, which increased activity by nearly 2-fold. Most of SP had lower efficiency than the native SP. Moreover, there were 6 SPs (SP1, SP7, SP8, SP14, SP17 and SP23) having no secretion with CelA (Figure 5.5a, b).

From the screening results of GH9/CelA (Figure 5.6), two results were almost consistent. Most SPs had higher secretion efficiency than the native SP (Figure 5.6b). SP4, SP7, SP13 and SP18 all had high secretion efficiency, increasing activity by more than 3-fold compared to the positive control. SP18, SP of beta-N-acetylhexosaminidase (from GEOTH_RS12070), showed

best efficiency for GH9/CelA secretion. SP23 had very low efficiency while SP15 showed no secretion on CMC plate (Figure 5.6a).

Secretion efficiency of SPs on GH48/CelA was not tested on CMC plate, because of low activity. From the result of enzyme assay (Figure 5.7), most SPs had similar efficiency with SP of CelA. Only SP16, SP of a hypothetical protein (from GEOTH_RS06780), had a promising result, which increased activity by more than 6-fold compared to the positive control.

In addition, the synergistic relationship between GH9/CelA and GH48/CelA on the degradation of cellulose was studied by using PASC as the substrate in different concentrations (0.4 mg/mL and 10mg/mL). The degrees of synergy (DOS) was calculated as described in material and methods. From the synergy assay, the greatest DOS was obtained with a volume ratio of 1: 1 (GH9/CelA: GH9/CelA). Since the two secreted enzyme in the supernatant was concentrated before enzyme reaction (described in material and methods), the actual volume ratio of supernatant for GH9/CelAand GH48/CelA was 1:2.

5.2.5 Comparison of SPs for different protein secretion

From the above screening results of 25 SPs for different protein secretion, no correlation was found between secretion efficiency and D scores. It means the efficiency depends on the signal peptide but not the D score.

To compare secretion efficiency of SPs for different protein, the top 10 SPs for AmyE, EglS and CelA secretion were respectively picked up and listed (Figure 5.9a). The comparison of them resulted in three common SPs, SP4, SP9 and SP21 (Figure 5.9b). SP9, SP of beta-lactamase (from GEOTH_RS12075), had relatively high secretion efficiency for all of AmyE, EglS and CelA in *G. thermoglucosidasius* 95A1. SP21 had the highest efficiency for both of AmyE and EglS, but not very efficient for CelA secretion.

5.3 Discussion

Here, 25 SPs were screened and characterized for their secretion efficiency for different heterologous protein. These SPs were predicted by the online SignalP 4.1 Server and ranked according to their D scores calculated by the server. From the above results, it can conclude that secretion efficiency of a SP is not correlated to its D score. D score shows the probability of an amino acid sequence being a signal peptide, but cannot indicate its secretion efficiency. The conclusion was also verified by other studies [12,13]. However, SignalP is still a very useful tool for SP prediction, which can be used for finding more SPs for screening.

A new cloning strategy (Figure 5.1a) was used for construction of 25 SP with different target genes, but it can be used for more SPs. Although this strategy added a more step of TA cloning (clone PCR product into T vector), it is more convenient for SP sequencing and can also increase the cloning efficiency. Direct cloning was also tried at the beginning of this study, but mutations or dimers were found in sequencing. The correct constructs might give a burden for *E. coli* growth. This problem was avoided by using TA cloning. Almost all of the colonies from TA cloning were correct. Moreover, Golden Gate assembly was employed for the next step of cloning, which kept the strategy efficient. Therefore, this strategy can also be applied for SP library construction. The only concern is that the transformation efficiency of *G. thermoglucosidasius* might be a limited factor. Electroporation was used for plasmid transformation into the strain, the efficiency of which is related to the size of plasmid. When transform the CelA secretion plasmids into *G. thermoglucosidasius* 95A1, the efficiency was very low (≈ 40 transformants/ μg DNA).

The efficiency of different SPs was screened by plate assay or enzyme assay through measuring reduced sugar. Both methods can be used for high throughput screening, but plate

assay is more convenient and enzyme assay is more accurate. Therefore, plate assay can be used for first-round screening while enzyme assay can be carried out for rescreening. In screening of this study, the results from two assays were almost consistent except ones for CelA. It might be the SPs changed substrate specificity. CelA with different SPs was tested on CMC plates and enzyme assay using PASC as the substrate. Similar result was also observed in another study of screening SPs for secretion of exoglucanase CelK from *Clostridium thermocellum* in *B. subtilis*. It tested CelK with SPs on different substrates and found activities of SP combined with CelK varying on different substrates suggested that SP might influence substrate specificity [13].

In previous work by Dr Angel Rivera, different strains of *G. thermoglucosidasius* were studied for their ability of native protein secretion on different substrates. It was found total protein secreted by *G. thermoglucosidasius* C56-Y93 was more than *G. thermoglucosidasius* 95A1. Thus, C56 was tried as the host firstly for amylase secretion and it indeed secreted more protein from the assay of starch plate. However, the phenotype was not stable possibly because of instability of the plasmid. In addition, C56 was not easily genetic engineered in our previous experience (Chapter 3). Therefore, *G. thermoglucosidasius* 95A1 was chosen as the host for protein secretion, since it is well studied and has developed genetic tools. Besides, it is a good producer of biofuels and chemicals, which makes it a promising strain for consolidated bioprocessing.

There are a lot of studies about screening different SPs for protein secretion in *Bacillus*. For *Geobacillus*, this is the first time of screening SPs for optimizing protein secretion. The strategy can be employed for screening all natural SPs from this strain or other *Geobacillus* species. The similarity of 25 SPs from *G. thermoglucosidasius* C56-Y93 with ones from *G.*

thermoglucosidasius 95A1 is very high except SP4 from endo-1, 4-beta-xylanase. 95A1 doesn't have any xylanase genes.

By comparison of top ten efficient SPs for different protein secretion, one SP (SP9) was found to fit all of AmyE, EglS and CelA. However, it cannot infer this SP can fit other secretory targets. In screening results of GH9/CelA and GH48/CelA, its secretion efficiency was low. As other studies, there is no one-size-fits-all SP. For a better result of protein secretion, different SPs need to be screened. Secretion efficiency is considered to be related to net charge, hydrophobicity and the length of the SP [20,21]. Moreover, it is related to the N-terminal part of the mature protein [13]. According to the results of CelA and GH48/CelA, secretion efficiency was also related to the whole mature protein but not only the N-terminal part.

5.4 Conclusion

In conclusion, a strategy was established for the optimization of protein secretion in *G. thermoglucosidasius* 95A1 by screening 25 SPs from *G. thermoglucosidasius* C56-Y93. Different heterologous thermostable enzymes were chosen as the secretion targets, including α-amylase AmyE from *G. stearothermophilus*, endoglucanase EglS from *B. subtilis* and bifunctional cellulase CelA from *C. bescii*. Optimal SPs were found for the different secretory proteins, which increased the secretion compared to their native SPs. One SP was found to fit all the three proteins but it is possible to be inefficient for other secretory proteins. Since amylase and cellulase are critical hydrolase in biofuel production, *G. thermoglucosidasius* is promising strain for consolidated bioprocessing. Besides, the chapter established this thermophile as a good candidate for thermostable proteins producer.

5.5 Material and methods

5.5.1 Plasmid construction

All plasmids used in this study are listed in Table 5.2. All primers used in this study are listed in Table 5.3.

The *amyE* secretion plasmids with different SPs were constructed as follows. The *amyE* receiving plasmid of pJZ57 was constructed firstly by PCR amplifying the *amyE* promoter, the truncated gene (without its SP) and the terminator from *G. stearothermophilus* 10 genomic DNA using the respective primer pairs of JZ327F/JZ328R, JZ329F/JZ197R and JZ333F2/ JZ333R2. The three DNA fragments were then assembled into the EcoRI and SphI sites of pJZ04n by using Gibson Assembly. To screen different SPs, small DNA fragments encoding 25 SPs with BsaI sites on ends were PCR amplified from *G. thermoglucosidasius* C56-YS93 by using respective primers pairs (from JZ601F to JZ650R in Table 5.2) and then cloned into pGEM-T easy Vector (Promega), respectively. Those SPs fragments on T-vectors were sequenced and then inserted into pJZ57 by the cloning method of Golden Gate for screening.

The cellulase secretion plasmids with different SPs were constructed by the same strategy. The 25 T-Vectors with different small DNA fragments of SPs were reconstructed by using relevant primers pairs (from JZ695F to JZ744R in Table 5.2). These new constructs of SPs can be used for fusing with different protein for secretion. The *eglS* receiving plasmid of pJZ81 was constructed firstly. The *ldh* promoter was PCR amplified from *G. thermodenitrificansius* using the primers of JZ413F2 and JZ414R2 , the truncated *eglS* gene (without its SP) was amplified from *B. subtilis* 168 using the primers of JZ524F2 and JZ525R2, and the *recA* terminator was amplified from *G. thermoglucosidasius* 95A1 using the primers of JZ526F2 and JZ527R2. The three DNA fragments were then assembled into the SacI and SphI sites of pJZ04n by Gibson

Assembly. The *celA* receiving plasmid of pJZ83 was constructed by replacing the *eglS* gene in pJZ81 with the truncated *celA* gene from *Caldicellulosiruptor bescii*. Briefly, the *celA* gene was PCR amplified by using the primers of JZ532F and JZ530R and then cloning the resulting DNA fragment into the XhoI and AatII sites of pJZ81. The GH9/*celA* and GH48/*celA* receving plasmids of pJZ85 and pJZ87 were similarly constructed using the primer pairs JZ532F/JZ531R and JZ533F/JZ530R to amplify the GH9/*celA* and GH48/*celA*, respectively.

5.5.2 Enzyme assays

The supernatant of culture was used for determination of secreted enzyme activity. The culture was grown overnight in TGP medium at 55°C, 250 rpm. The supernatant of culture was collected and passed through 0.22 μm polyethersulfone syringe filter before adding to the enzyme reaction. The α-Amylase and cellulase activity was determined by measuring the amount of released reducing sugar by DNS method (described in 2.3.2.1) or pHBAH assay (described in 2.3.2.2).

5.5.3.1 α-Amylase activity assay

The α-amylase activity was determined by DNS method. The enzymatic reaction (40 μL) was performed in 10 mM Tris-HCl buffer (pH 7.0) containing 1% starch and 2.5 μL supernatant of culture. The reaction was incubated for 20 minutes at 70°C on a thermal cycler.

5.5.3.2 Cellulase activity assay

The endoglucanase (EglS) activity was determined by DNS method. The enzymatic reaction (40 μL) was performed in 50 mM citric acid buffer (pH 6.0) with 1% CMC and 2 μL supernatant of culture. The reaction was incubated for 30 minutes at 60°C on a thermal cycler.

The CelA and GH9/CelA activity were determined by pHBAH method. The GH48/CelA activity was determined by measuring cellobiose in enzyme reaction using HPLC. The enzymatic reaction (1 mL) was performed in CelA buffer (50 mM sodium-citrate, 150 mM NaCl, pH 6.0) with 1% PASC (preparation described in 2.3.3) and 0.1 mL supernatant of culture. The reaction was incubated for 24 h at 75°C by shaking.

5.5.3 Determination of the degree of the synergy

GH9/SP14 and GH48/SP16 were cultured in TMLM (55°C, 250 rpm, 16h) and the supernatant was concentrated to about 1/4 and 1/8 volume, respectively, by using Amicon Ultra 15 mL Centrifugal Filters. Different concentrations of PASC were used for enzyme reaction as indicated. The degrees of synergy (DOS) were calculated as follows:

DOS = (Glucose equivalents released by enzyme combination)/(Sum of glucose equivalents released by individual enzymes) [15]

5.6 Tables

Table 5.1 25 Signal peptides from *G. thermoglucosidasius* C56-YS93

No.	Gene	Protein name	Amino acid sequence	Length	D[a] score
SP1	GEOTH_RS07945	Sporulation protein	MNKKIVFSLAASLAIVGASFTAKA	24	0.861
SP2	GEOTH_RS01540	Serine protease	MKKWKKTAVSLGLASALVLPSFAQA	25	0.849
SP3	GEOTH_RS01725	S-layer protein	MKRTFLHIALSLLAAMLALPAMNASA	26	0.829
SP4	GEOTH_RS17105	Peptidase P60	MKQFVTLVSLSFLVVFSSLFAHTSSAEA	28	0.815
SP5	GEOTH_RS11065	Endo-1,4-beta-xylanase	MRNVLRKPIVAGLSLTLLLPIGVNDTSA	28	0.800
SP6	GEOTH_RS19995	S-layer protein	MAYQPKSYRKFLAGSVSAALVATAVGPVVA NA	32	0.799
SP7	GEOTH_RS01780	Mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase	MRIGVQIRKFAALLSVLILLVSYAISPAYA	30	0.799
SP8	GEOTH_RS02345	Hypothetical protein	LKKLLLSITSSFFLAFGFSGAASA	24	0.781
SP9	GEOTH_RS12075	Beta-lactamase	MSNRFVSVVLLSVMLSSAIFFSPPSVLA	28	0.780
SP10	GEOTH_RS06855	Serine-type D-Ala-D-Ala carboxypeptidase	MKLWKLIVLFIVAVAMLFSCIPDQAKA	27	0.757
SP11	GEOTH_RS03295	Nuclease	MKKFVSALAIIVSTAIFPGNSFA	23	0.753
SP12	GEOTH_RS06580	ErfK/YbiS/YcfS/YnhG family protein	MRWILAAMLVLSSFFSISASAA	22	0.750
SP13	GEOTH_RS14050	Hypothetical protein	MNKTKSYLSFLLSFVLVLSTLGGAGIAQA	29	0.741
SP14	GEOTH_RS17840	Hypothetical protein	MFKKGYLSILSLVMGFTFFSTNTFA	25	0.738
SP15	GEOTH_RS08035	ABC transporter substrate-binding protein	MFKSKLSFLITAILTLVIILAGCGKNEKA	29	0.734
SP16	GEOTH_RS06780	Hypothetical protein	VKRMLTGCLLASLLFAFPAMA	21	0.723
SP17	GEOTH_RS05045	Ig domain-containing protein	VAEKRKFLWLLMALLLCVAFGNVPAVAFG	29	0.719
SP18	GEOTH_RS12070	Beta-N-acetylhexosaminidase	MLSFYKKITVILVAVVMLFVPWTSPQA	27	0.719
SP19	GEOTH_RS12265	Hypothetical protein	MKKWILAMLSVSVLCLLVYFVIGQAENVFA	30	0.716
SP20	GEOTH_RS03735	Hypothetical protein	MKRKPWKVMTAAALTSSLLLASACTSSG	28	0.712
SP21	GEOTH_RS18390	Hypothetical protein	MVLKKILSGVLGLSLLLGGTNFAFA	25	0.710
SP22	GEOTH_RS06345	Stage III sporulation protein AE	LKVKAVAVVGLFFFFFSPFVVQA	23	0.706
SP23	GEOTH_RS10570	Hypothetical protein	VKLPKWLRKVLVVTITVCTFGLVTPPASLMA	31	0.706

Table 5.1 (cont.)

No.	Gene	Protein name	Amino acid sequence	Length	D[a] score
SP24	GEOTH_RS09445	NLP/P60 protein	MKKSFILTGTIISSLLADQTAFA	23	0.703
SP25	GEOTH_RS01075	ABC transporter substrate-binding protein	MMKRKWLYFSLIALLILILTACGAKQSSA	29	0.699

[a] Discrimination score.

Table 5.2 Plasmids and strains used in this study

Plasmid or strain	Characteristics	Source or reference
Plasmids		
pJZ04n	Cm^R; *E. coli-Bacillus* shuttle vector; pNW33N derivative	The 4th chapter
pJZ04	Spc^R; *E. coli-Bacillus* shuttle vector	The 3rd chapter
pJZ57	pJZ04n containing truncated *amyE* (without its SP) from *G. stearothermophilus* 10 with its promoter and terminator	This study
pJZ81	pJZ04n containing truncated *eglS* (without its SP) from *B. subtilis* 168 with promoter *ldh* from *G. thermodenitrificansius* and terminator of *recA* from *G. thermoglucosidasius* 95A1	This study
pJZ83	*celA* from *Caldicellulosiruptor bescii* replaces of *eglS* in pJZ81	This study
pJZ85	GH9 of *celA* from *C. bescii* replaces of *eglS* in pJZ81	This study
pJZ87	GH48 of *celA* from *C. bescii* replaces of *eglS* in pJZ81	This study
Strains		
Geobacillus thermoglucosidasius BGSC 95A1	Wild type, $DSM2452^T$	BGSC
G. thermoglucosidasius C56-YS93	Source of SPs	
G. stearothermophilus 10	Source of *amyE*, its promoter and terminator	BGSC
G. thermodenitrificans 94A1	Source of P_{ldh}	BGSC
Bacillus subtilis 168	Source of *eglS*	BGSC

Table 5.3 Primers used in this study

Primer	Sequence	Plasmid
JZ327F	TGAGTAAACTTGGTCTGACAGTTAGAATTCTTTGCACAAAATGGTTGATG	pJZ57
JZ328R	CGGAGACCCCTTTATGGTGGTCTCGAATGCCCTTCCCCCTTAATCAAATG	
JZ329F	CGAGACCACCATAAAGGGGTCTCCGCACCGTTTAACGGCACCATGATGC	
JZ197R	CATATGACCGGTTTATCAAGGCCATGCCACCAACC	
JZ333F2	ATGGCCTTGATAAACCGGTCATATGTGCCTGCGATCGCGTTGTAAAG	
JZ333R2	TATGACCATGATTACGCCAAGCTTGCATGCATCTATGGAAGCGGAC	
JZ413F2	TGGTCTGACAGTTAGAATTCGAGCTCTTGGCGTCTATCCGGCTTCG	pJZ81
JZ414R2	TGGTGGTCTCACTCGAGACTCCTCCTTTCTAGATGTGAATACATTCACAAATTATTTTAC	
JZ524F2	AAGGAGGAGTCTCGAGTGAGACCACCATAAAGGGGTCTCAGGCAGGGACAAAAACGCCAGTAG	
JZ525R2	TCTCTCGCACCGGTAAGACGTCTTACTAATTTGGTTCTGTTCCCCAAATC	
JZ526F2	CAAATTAGTAAGACGTCTTACCGGTGCGAGAGAGGGACTTTGTTTTTC	
JZ527R2	AACAGCTATGACCATGATTACGCCAAGCTTGCATGCACGCGCCCCCATTTTG	
JZ530R	GC GACGTC TTATTATTGATTGCCAAACAGTATTTCATATG	pJZ83, pJZ87
JZ532F	ATCTCGAGTGAGACCACCATAAAGGGGTCTCAGGCAGGTTCGTTTAACTATGGGGAAGC	pJZ83, pJZ85
JZ531R	GCGACGTCTTATTATACCTTTATCTGTCCACCTGCTACAG	pJZ85
JZ533F	GCCTCGAGTGAGACCACCATAAAGGGGTCTCAGGCATTTGTTGAAGCTGGTATAAATG	pJZ87
JZ645F	GCGGTCTCACGAGATGAACAAAAAAATCGTATTTTCC	SP1 for *amyE*
JZ646R	ATGGTCTCGTGCCGCTTTCGCTGTAAACGACGCAC	
JZ647F	GCGGTCTCACGAGATGAAAAAATGGAAAAAAACAG	SP2 for *amyE*
JZ648R	ATGGTCTCGTGCCGCCTGGGCAAAGCTAGGCAAG	
JZ649F	GCGGTCTCACGAGATGAAACGTACATTTCTTC	SP3 for *amyE*
JZ650R	ATGGTCTCGTGCCGCGGAAGCATTCATGGCAG	
JZ601F	GCGGTCTCACGAGATGAAACAATTTGTCACACTAG	SP4 for *amyE*
JZ602R	ATGGTCTCGTGCCGCTTCTGCGCTAGAAGTATG	
JZ603F	GCGGTCTCACGAGATGCGGAACGTTTTACGCAAAC	SP5 for *amyE*
JZ604R	ATGGTCTCGTGCCGCAGATGTGTCATTCACTC	

Table 5.3 (cont.)

Primer	Sequence	Plasmid
JZ605F	GCGGTCTCACGAGATGGCTTACCAACCAAAGTC	SP6 for *amyE*
JZ606R	ATGGTCTCGTGCCGCGTTGGCTACAACTGGAC	
JZ607F	GCGGTCTCACGAGATGCGGATAGGAGTACAAATAAG	SP7 for *amyE*
JZ608R	ATGGTCTCGTGCCGCATATGCCGGAGAAATGG	
JZ609F	GCGGTCTCACGAGATGAAAAAATTACTATTATC	SP8 for *amyE*
JZ610R	ATGGTCTCGTGCCGCCGATGCAGCTCCTGAAAATC	
JZ611F	GCGGTCTCACGAGATGAGCAATCGGTTTGTTTCC	SP9 for *amyE*
JZ612R	ATGGTCTCGGTGCCGCTAAAACGGATGGAGGCG	
JZ613F	GCGGTCTCACGAGATGAAGCTCTGGAAACTAATTG	SP10 for *amyE*
JZ614R	ATGGTCTCGTGCCGCCTTCGCTTGATCCGGAATAC	
JZ615F	GCGGTCTCACGAGATGAAGAAATTCGTTAGCGC	SP11 for *amyE*
JZ616R	ATGGTCTCGTGCCGCAAAACTATTTCCCG	
JZ617F	ATGGTCTCACGAGATGCGATGGATATTGGCAGC	SP12 for *amyE*
JZ618R	ATGGTCTCGTGCCGCCGCGGAAGCAGAAATGG	
JZ619F	GCGGTCTCACGAGATGAACAAAACAAAAAGTTATTTATC	SP13 for *amyE*
JZ620R	ATGGTCTCGTGCCGCTTGGGCAATTCCTGCACCAC	
JZ621F	GCGGTCTCACGAGATGTTCAAAAAAGGTTATTTATC	SP14 for *amyE*
JZ622R	GCGGTCTCGTGCCGCAAAAGTATTAGTTGAG	
JZ623F	GCGGTCTCACGAGATGTTCAAATCAAAACTTTC	SP15 for *amyE*
JZ624R	ATGGTCTCGTGCCGCTTTCTCATTTTTGCCGC	
JZ625F	TGGGTCTCACGAGATGAAACGCATGTTGACAG	SP16 for *amyE*
JZ626R	ATGGTCTCGTGCCGCCATAGCGGGAAAGGC	
JZ627F	GCGGTCTCACGAGATGGCGGAGAAGAGAAAATTTTTATG	SP17 for *amyE*
JZ628R	ATGGTCTCGTGCCCCAAACGCTACAGCAGGAAC	

Table 5.3 (cont.)

Primer	Sequence	Plasmid
JZ629F	GCGGTCTCACGAGATGTTGTCATTTTATAAAAAAATAAC	SP18 for *amyE*
JZ630R	ATGGTCTCGTGCCGCTTGCGGCGAAGTCCATG	
JZ631F	GCGGTCTCACGAGATGAAAAAATGGATTTTGGC	SP19 for *amyE*
JZ632R	ATGGTCTCGTGCCGCAAAAACATTCTCCGCTTG	
JZ633F	GCGGTCTCACGAGATGAAACGAAAACCGTGGAAAG	SP20 for *amyE*
JZ634R	ATGGTCTCGTGCCCCACTGGAAGTGCAAGCGG	
JZ635F	GCGGTCTCACGAGATGGTATTGAAAAAAATC	SP21 for *amyE*
JZ636R	ATGGTCTCGTGCCGCAAAGGCAAAGTTTGTTC	
JZ637F	GCGGTCTCACGAGATGAAAGTAAAGGCAGTAGC	SP22 for *amyE*
JZ638R	ATGGTCTCGTGCCGCTTGTACCACAAATGGAG	
JZ639F	GCGGTCTCACGAGATGAAACTGCCAAAATGGTTG	SP23 for *amyE*
JZ640R	ATGGTCTCGTGCCGCCATTAATGAAGCCGGCG	
JZ641F	GCGGTCTCACGAGATGAAAAAATCATTTATTCTG	SP24 for *amyE*
JZ642R	ATGGTCTCGTGCCGCGAAAGCAGTTTGGTCTG	
JZ643F	GCGGTCTCACGAGATGATGAAAAGGAAATGGC	SP25 for *amyE*
JZ644R	ATGGTCTCGTGCCGCAGATGATTGTTTCGCG	
JZ695F	GCGGTCTCACGAGATGAACAAAAAAATCGTATTTTCC	SP1
JZ696R	ATGGTCTCGTGCCGCTTTCGCTGTAAACGACGCAC	
JZ697F	GCGGTCTCACGAGATGAAAAAATGGAAAAAAACAG	SP2
JZ698R	ATGGTCTCGTGCCGCCTGGGCAAAGCTAGGCAAG	
JZ699F	GCGGTCTCACGAGATGAAACGTACATTTCTTC	SP3
JZ700R	ATGGTCTCGTGCCGCGGAAGCATTCATGGCAG	
JZ701F	GCGGTCTCACGAGATGAAACAATTTGTCACACTAG	SP4
JZ702R	ATGGTCTCGTGCCGCTTCTGCGCTAGAAGTATG	

Table 5.3 (cont.)

Primer	Sequence	Plasmid
JZ703F	GCGGTCTCACGAGATGCGGAACGTTTTACGCAAAC	SP5
JZ704R	ATGGTCTCGTGCCGCAGATGTGTCATTCACTC	
JZ705F	GCGGTCTCACGAGATGGCTTACCAACCAAAGTC	SP6
JZ706R	ATGGTCTCGTGCCGCGTTGGCTACAACTGGAC	
JZ707F	GCGGTCTCACGAGATGCGGATAGGAGTACAAATAAG	SP7
JZ708R	ATGGTCTCGTGCCGCATATGCCGGAGAAATGG	
JZ709F	GCGGTCTCACGAGATGAAAAAATTACTATTATC	SP8
JZ710R	ATGGTCTCGTGCCGCCGATGCAGCTCCTGAAAATC	
JZ711F	GCGGTCTCACGAGATGAGCAATCGGTTTGTTTCC	SP9
JZ712R	ATGGTCTCGGTGCCGCTAAAACGGATGGAGGCG	
JZ713F	GCGGTCTCACGAGATGAAGCTCTGGAAACTAATTG	SP10
JZ714R	ATGGTCTCGTGCCGCCTTCGCTTGATCCGGAATAC	
JZ715F	GCGGTCTCACGAGATGAAGAAATTCGTTAGCGC	SP11
JZ716R	ATGGTCTCGTGCCGCAAAACTATTTCCCG	
JZ717F	ATGGTCTCACGAGATGCGATGGATATTGGCAGC	SP12
JZ718R	ATGGTCTCGTGCCGCCGCGGAAGCAGAAATGG	
JZ719F	GCGGTCTCACGAGATGAACAAAACAAAAAGTTATTTATC	SP13
JZ720R	ATGGTCTCGTGCCGCTTGGGCAATTCCTGCACCAC	
JZ721F	GCGGTCTCACGAGATGTTCAAAAAAGGTTATTTATC	SP14
JZ722R	GCGGTCTCGTGCCGCAAAAGTATTAGTTGAG	
JZ723F	GCGGTCTCACGAGATGTTCAAATCAAAACTTTC	SP15
JZ724R	ATGGTCTCGTGCCGCTTTCTCATTTTTGCCGC	
JZ725F	TGGGTCTCACGAGATGAAACGCATGTTGACAG	SP16
JZ726R	ATGGTCTCGTGCCGCCATAGCGGGAAAGGC	
JZ727F	GCGGTCTCACGAGATGGCGGAGAAGAGAAAATTTTTATG	SP17
JZ728R	ATGGTCTCGTGCCCCAAACGCTACAGCAGGAAC	
JZ729F	GCGGTCTCACGAGATGTTGTCATTTTATAAAAAAATAAC	SP18
JZ730R	ATGGTCTCGTGCCGCTTGCGGCGAAGTCCATG	

Table 5.3 (cont.)

Primer	Sequence	Plasmid
JZ731F	GCGGTCTCACGAGATGAAAAAATGGATTTTGGC	SP19
JZ732R	ATGGTCTCGTGCCGCAAAAACATTCTCCGCTTG	
JZ733F	GCGGTCTCACGAGATGAAACGAAAACCGTGGAAAG	SP20
JZ734R	ATGGTCTCGTGCCCCACTGGAAGTGCAAGCGG	
JZ735F	GCGGTCTCACGAGATGGTATTGAAAAAAATC	SP21
JZ736R	ATGGTCTCGTGCCGCAAAGGCAAAGTTTGTTC	
JZ737F	GCGGTCTCACGAGATGAAAGTAAAGGCAGTAGC	SP22
JZ738R	ATGGTCTCGTGCCGCTTGTACCACAAATGGAG	
JZ739F	GCGGTCTCACGAGATGAAACTGCCAAAATGGTTG	SP23
JZ740R	ATGGTCTCGTGCCGCCATTAATGAAGCCGGCG	
JZ741F	GCGGTCTCACGAGATGAAAAAATCATTTATTCTG	SP24
JZ742R	ATGGTCTCGTGCCGCGAAAGCAGTTTGGTCTG	
JZ743F	GCGGTCTCACGAGATGATGAAAAGGAAATGGC	SP25
JZ744R	ATGGTCTCGTGCCGCAGATGATTGTTTCGCG	

5.7 Figures

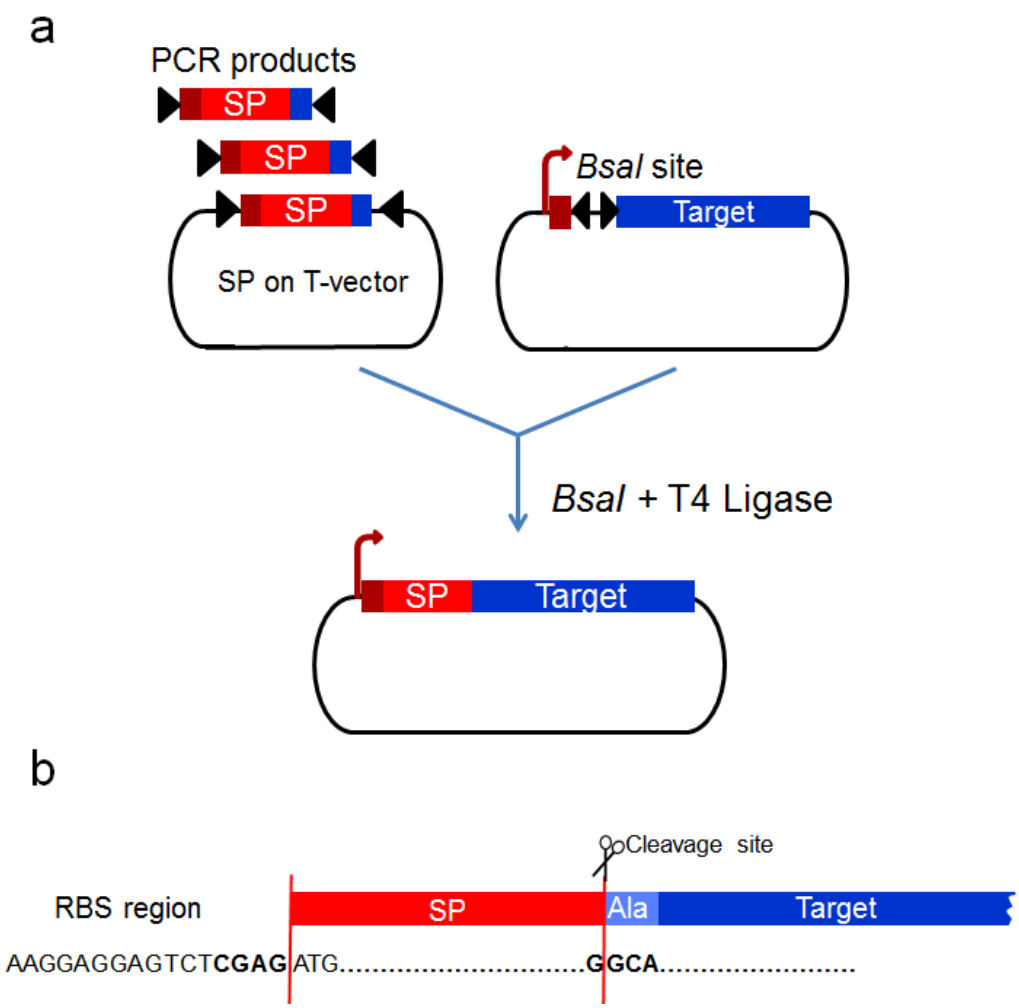

Figure 5.1 Cloning strategy of 25 signal peptides from *G. thermoglucosidasius* C56-YS93 fused with secretory protein targets. (a) Schematic overview of secretion plasmids construction. The PCR products were cloned into T vectors and then inserted into the receiving plasmid by Golden Gate assembly. (b) Sequence details of the secretory target protein fused with signal peptide (only for cellulase secretion plasmids in this chapter). The RBS was from *recA* of *G. thermoglucosidasius* 95A1 and an XhoI restriction sited was introduced right before the start codon. Ala was introduced as the first amino acid of the secretory protein. PCR products of the signal peptides were designed to have the same junctions (highlighted in bold) with the receiving plasmids. The junction at N-terminal was the last four nucleotide of XhoI. The other one at C-terminal was the designed last nucleotide of signal peptide "G", together with "GCA" for the first amino acid of secretory targets.

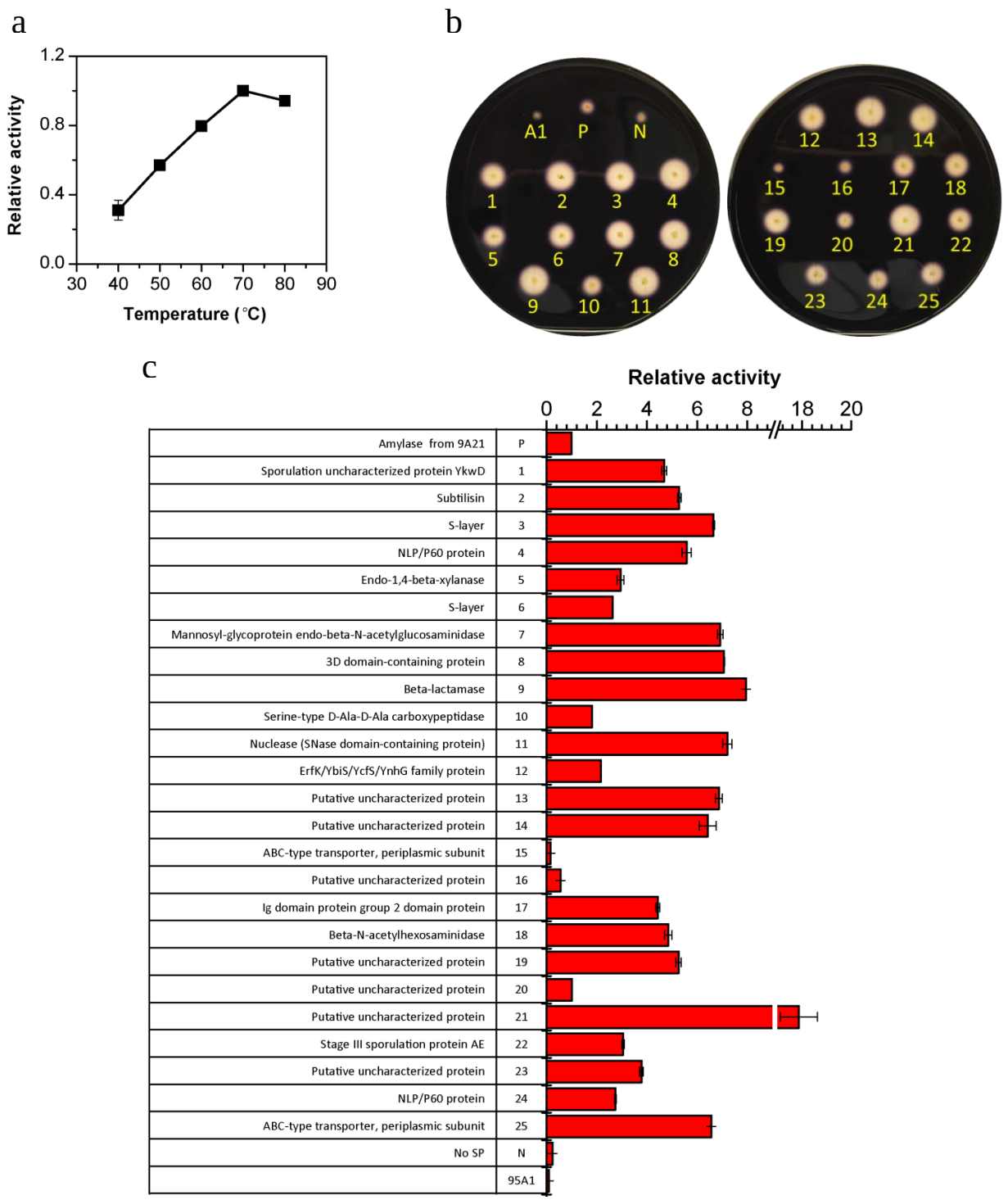

Figure 5.2 Comparison of 25 SPs used for heterologous α-amylase secretion in *G. thermoglucosidasius*. a, effect of temperature on activity; b, starch plates spotted with the liquid culture, and treated with iodine solution after overnight incubation at 55°C; c, relative enzyme activity in supernatant of culture by DNS assay. P, positive control, relative activity of AmyE with its native signal peptide was set as 1. N, negative control, no SP fused with protein. The figure shows the means and standard deviations obtained from three experiments.

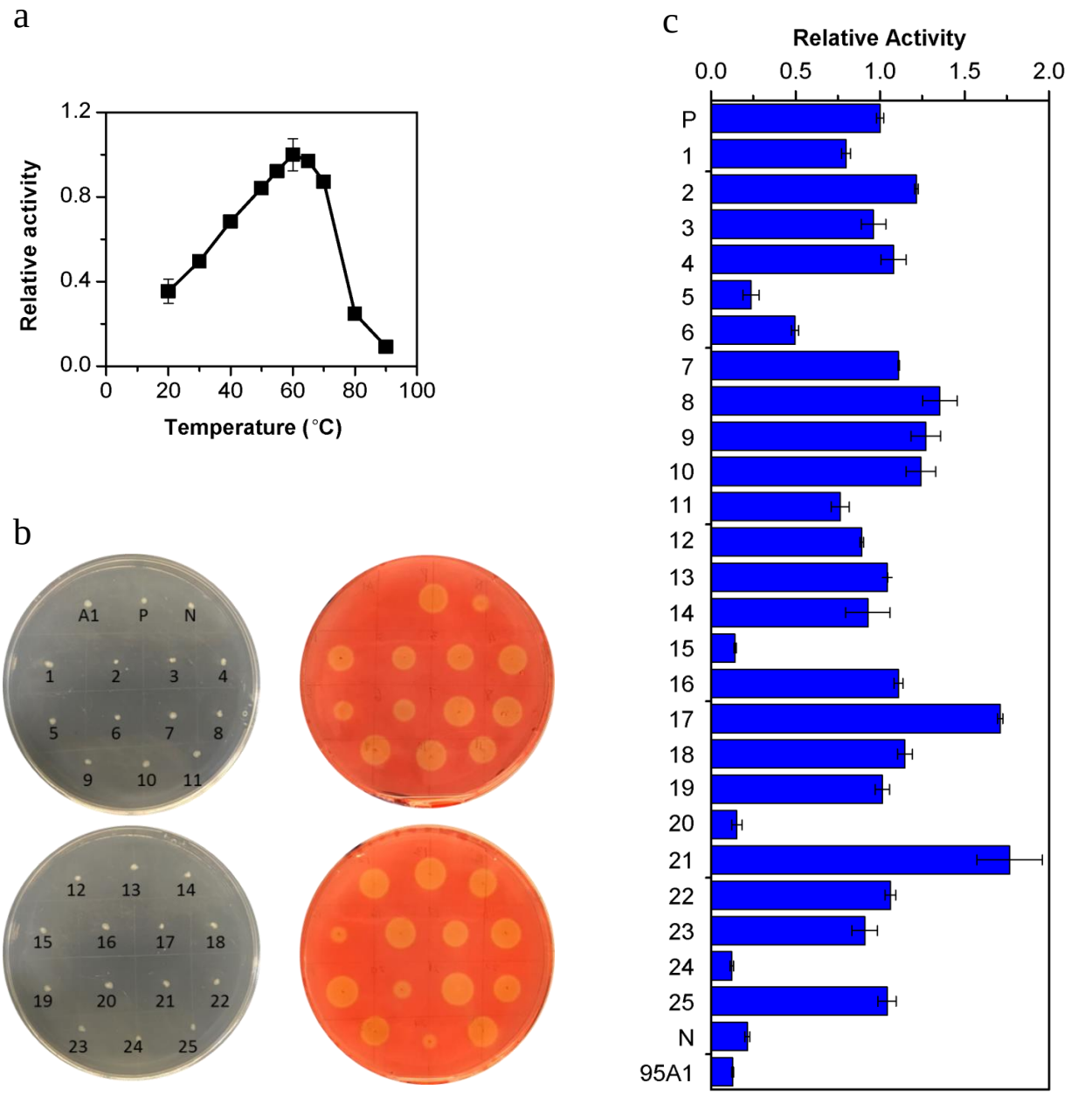

Figure 5.3 Comparison of 25 SPs used for heterologous EglS secretion in *G. thermoglucosidasius*. a, effect of temperature on activity; b, CMC plates spotted with the culture, and treated with Congo red solution after overnight incubation at 55°C; c, relative enzyme activity in supernatant of culture by DNS assay. P, positive control, relative activity of EglS with its native signal peptide was set as 1. N, negative control, no SP fused with protein. The figure shows the means and standard deviations obtained from three experiments.

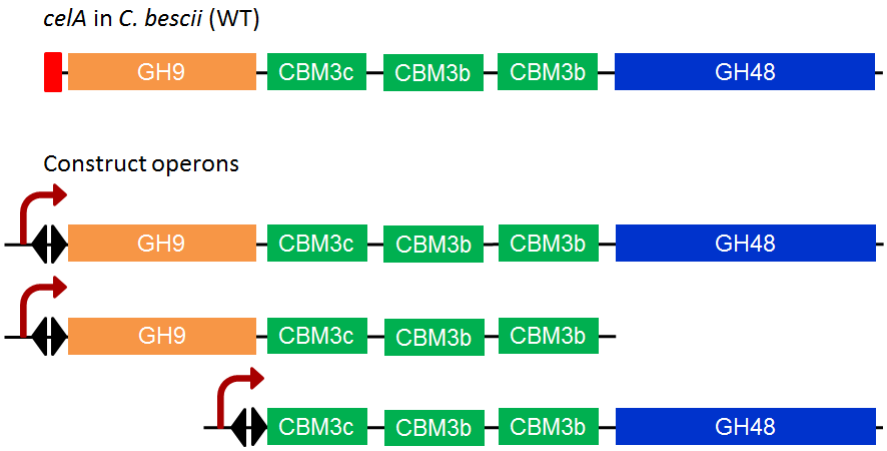

Figure 5.4 Schematic representation of *celA* (wild type), and its constructs for heterologous secretion in *G. thermoglucosidasius*. The red filled rectangle indicated its native signal peptide; The black filled triangle indicated BsaI site; GH9, glycoside hydrolase family 9 module; GH48, glycoside hydrolase family 8 module; CBM3b, CBM3c, different carbohydrate-binding modules.

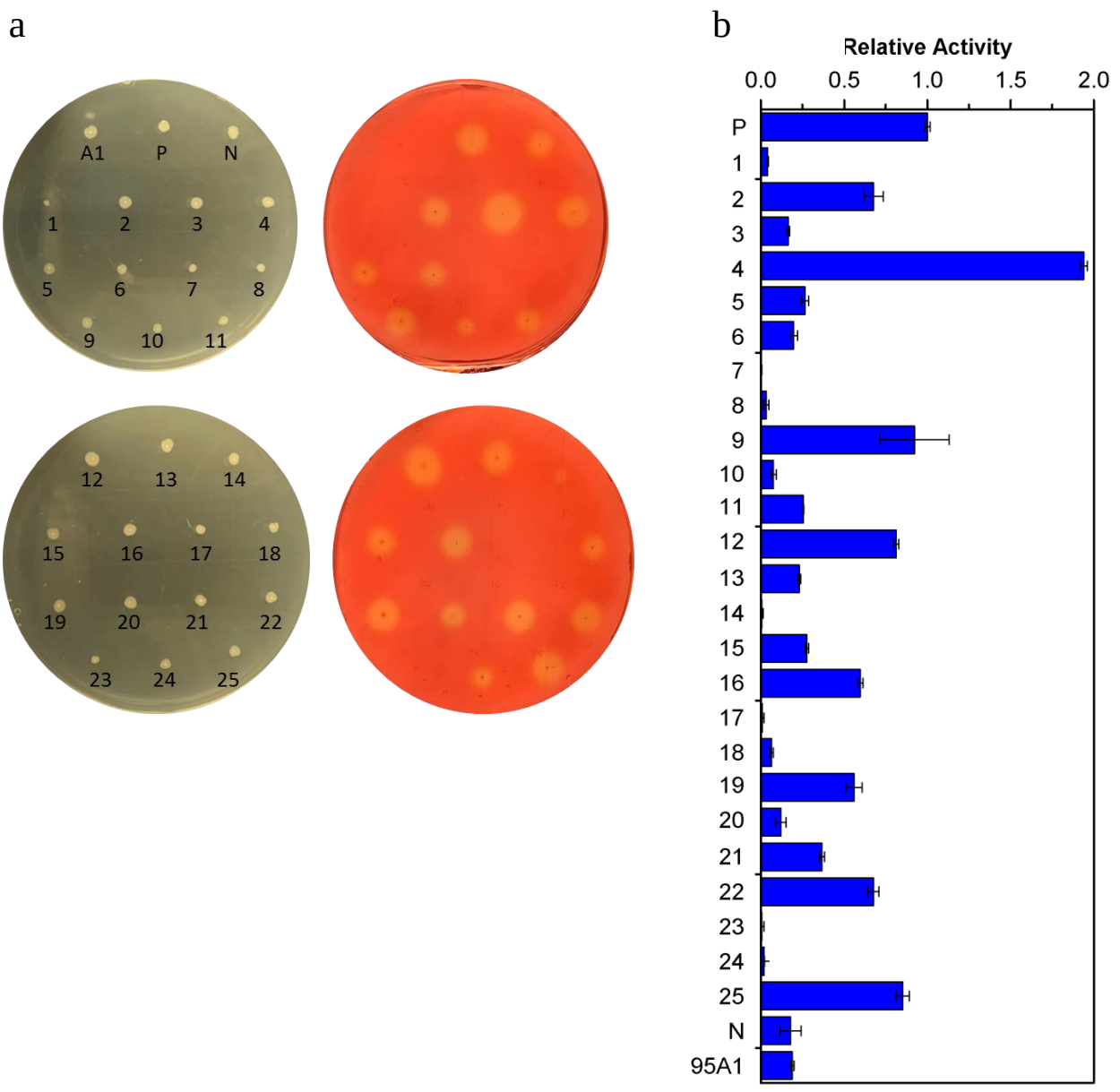

Figure 5.5 Comparison of 25 SPs used for heterologous CelA secretion in *G. thermoglucosidasius*. a, CMC plates spotted with the culture, and treated with Congo red solution after overnight incubation at 55°C; b, relative enzyme activity in supernatant of culture by pHBAH assay. P, positive control, relative activity of GH9/CelA with its native signal peptide was set as 1. N, negative control, no SP fused with protein. The figure shows the means and standard deviations obtained from three experiments.

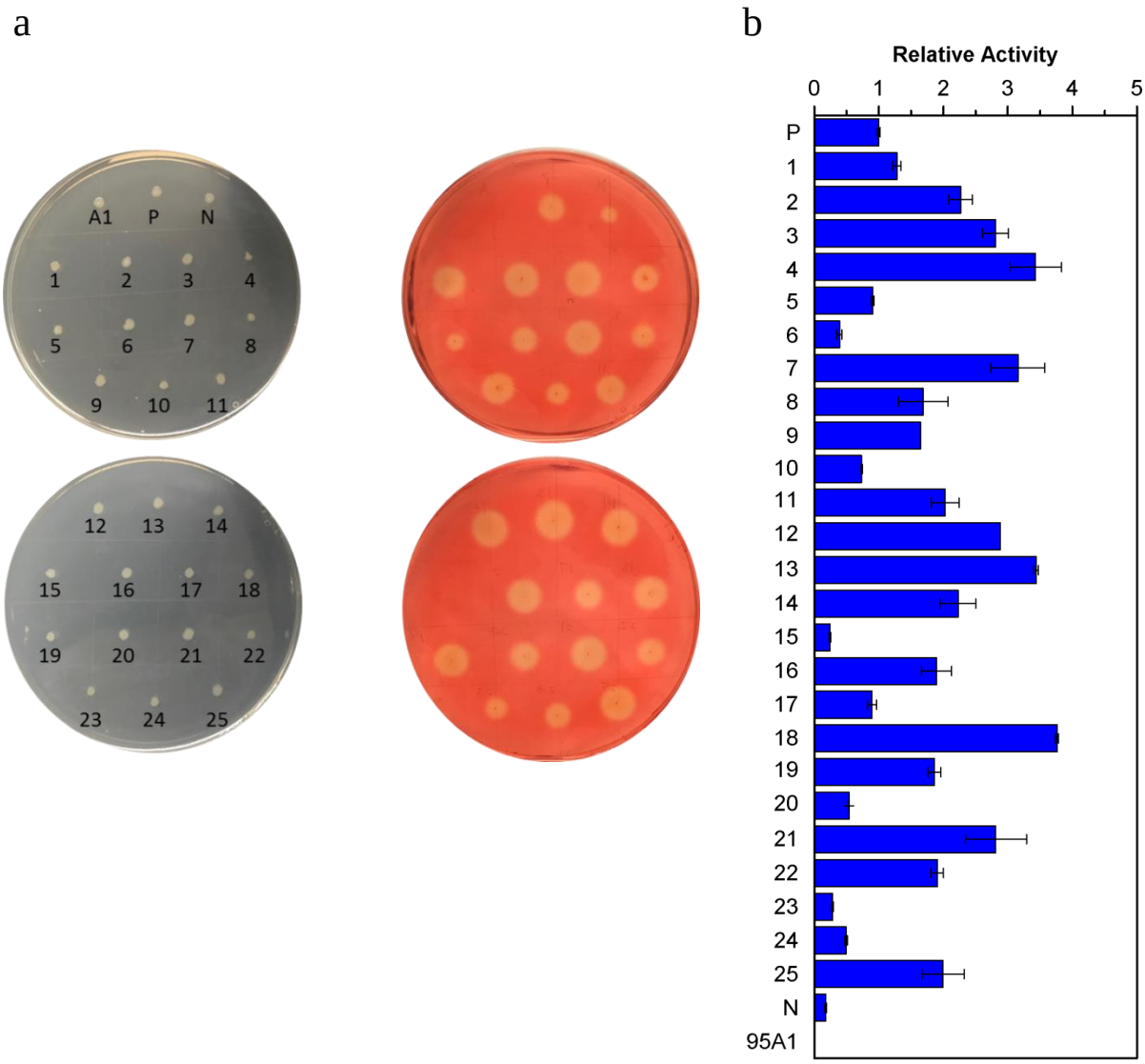

Figure 5.6 Comparison of 25 SPs used for heterologous GH9/CelA secretion in *G. thermoglucosidasius*. a, CMC plates spotted with the culture, and treated with Congo red solution after overnight incubation at 55°C; b, relative enzyme activity in supernatant of culture by pHBAH assay. P, positive control, relative activity of GH9/CelA with its native signal peptide was set as 1. N, negative control, no SP fused with protein. The figure shows the means and standard deviations obtained from three experiments.

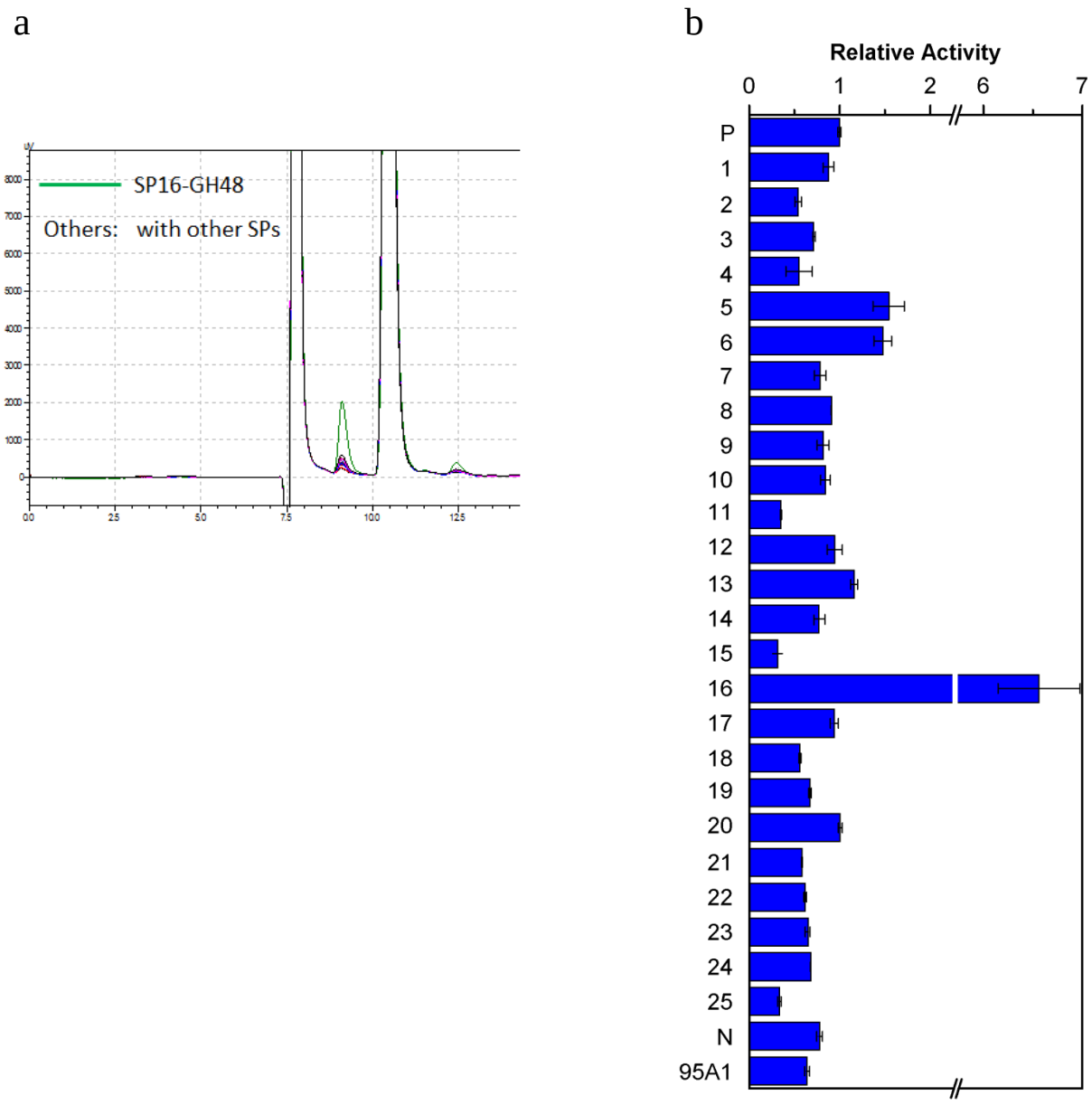

Figure 5.7 Comparison of 25 SPs used for heterologous GH48/CelA secretion in *G. thermoglucosidasius*. a, comparison of enzyme reaction profiles on HPLC. Cellobiose was analyzed by HPLC equipped with RI detector; b, relative enzyme activity of GH48/CelA with different SPs in supernatant of culture through measuring releasing cellobiose by HPLC. P, positive control, relative activity of GH48/CelA with CelA signal peptide was set as 1. N, negative control, no SP fused with protein. The figure shows the means and standard deviations obtained from three experiments.

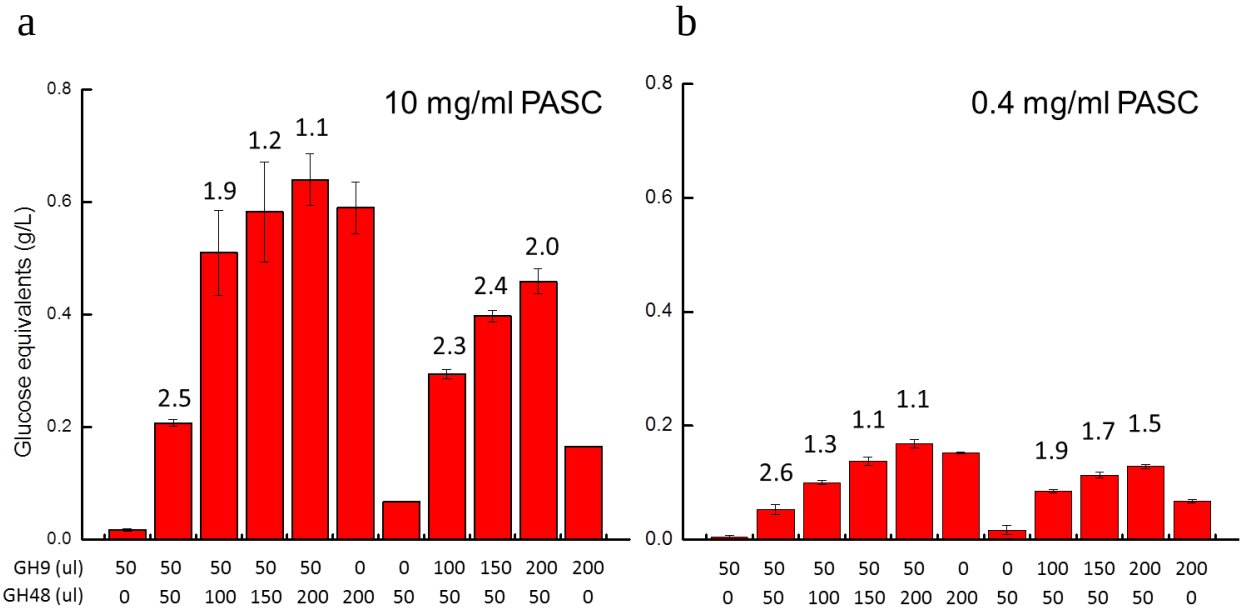

Figure 5.8 Synergistic relationship between GH9/CelA and GH9/CelA in hydrolysis of PASC. The experiments were carried out in CelA buffer at 75°C for 24 h. a, 10 mg/mL PASC were used as the substrate; b, 0.4 mg/mL PASC were used as the substrate. The numbers on top of the columns indicated the degrees of synergy calculated as described in Materials and Methods. The figure shows the means and standard deviations obtained from three experiments.

a

	AmyE	EglS	CelA
1	21	21	4
2	9	17	9
3	11	8	25
4	8	9	12
5	7	10	2
6	13	2	22
7	3	18	16
8	25	7	19
9	14	16	21
10	4	4	15

b

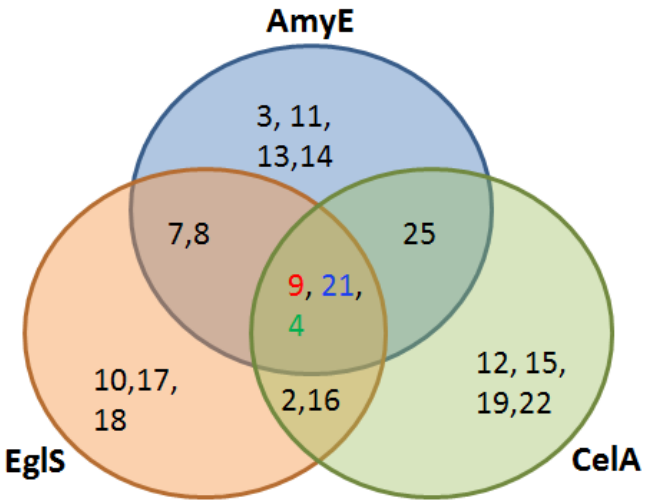

Figure 5.9 Comparison of top 10 SPs for different heterologous protein secretion in *G. thermoglucosidasius*. a, top 10 SPs for different protein were listed, which were ranked according to relative secretion efficiency; b, Comparison of top 10 SPs for different protein secretion.

5.8 References

1. Fong SS, Pogrebnyakov I, Jendresen CB, Nielsen AT (2017) Genetic toolbox for controlled expression of functional proteins in Geobacillus spp. PloS one 12: e0171313.
2. Ezeji TC, Bahl H (2006) Purification, characterization, and synergistic action of phytate-resistant α-amylase and α-glucosidase from Geobacillus thermodenitrificans HRO10. Journal of Biotechnology 125: 27-38.
3. Khemakhem B, Ali MB, Aghajari N, Juy M, Haser R, et al. (2009) Engineering of the α - amylase from Geobacillus stearothermophilus US100 for detergent incorporation. Biotechnology and Bioengineering 102: 380-389.
4. Li H, Zhang X (2005) Characterization of thermostable lipase from thermophilic Geobacillus sp. TW1. Protein expression and purification 42: 153-159.
5. Soliman NA, Knoll M, Abdel-Fattah YR, Schmid RD, Lange S (2007) Molecular cloning and characterization of thermostable esterase and lipase from Geobacillus thermoleovorans YN isolated from desert soil in Egypt. Process Biochemistry 42: 1090-1100.
6. Zhu W, Cha D, Cheng G, Peng Q, Shen P (2007) Purification and characterization of a thermostable protease from a newly isolated Geobacillus sp. YMTC 1049. Enzyme and Microbial Technology 40: 1592-1597.
7. Sookkheo B, Sinchaikul S, Phutrakul S, Chen S-T (2000) Purification and characterization of the highly thermostable proteases from Bacillus stearothermophilus TLS33. Protein expression and purification 20: 142-151.
8. Bartosiak-Jentys J, Hussein AH, Lewis CJ, Leak DJ (2013) Modular system for assessment of glycosyl hydrolase secretion in Geobacillus thermoglucosidasius. Microbiology 159: 1267-1275.
9. Holland A (2017) Optimisation of feedstock utilisation by Geobacillus thermoglucosidasius: University of Bath.
10. Harwood CR, Cranenburgh R (2008) Bacillus protein secretion: an unfolding story. Trends in Microbiology 16: 73-79.
11. Caspers M, Brockmeier U, Degering C, Eggert T, Freudl R (2010) Improvement of Sec-dependent secretion of a heterologous model protein in Bacillus subtilis by saturation mutagenesis of the N-domain of the AmyE signal peptide. Applied Microbiology and Biotechnology 86: 1877-1885.
12. Degering C, Eggert T, Puls M, Bongaerts J, Evers S, et al. (2010) Optimization of Protease Secretion in Bacillus subtilis and Bacillus licheniformis by Screening of Homologous and Heterologous Signal Peptides. Applied and Environmental Microbiology 76: 6370-6376.
13. Brockmeier U, Caspers M, Freudl R, Jockwer A, Noll T, et al. (2006) Systematic Screening of All Signal Peptides from Bacillus subtilis: A Powerful Strategy in Optimizing Heterologous Protein Secretion in Gram-positive Bacteria. Journal of Molecular Biology 362: 393-402.
14. Lin L, Meng X, Liu P, Hong Y, Wu G, et al. (2009) Improved catalytic efficiency of Endo-β-1, 4-glucanase from Bacillus subtilis BME-15 by directed evolution. Applied Microbiology and Biotechnology 82: 671.
15. Yi Z, Su X, Revindran V, Mackie RI, Cann I (2013) Molecular and biochemical analyses of CbCel9A/Cel48A, a highly secreted multi-modular cellulase by Caldicellulosiruptor bescii during growth on crystalline cellulose. PloS one 8: e84172.
16. Chung D, Young J, Bomble YJ, Vander Wall TA, Groom J, et al. (2015) Homologous expression of the Caldicellulosiruptor bescii CelA reveals that the extracellular protein is glycosylated. PloS one 10: e0119508.
17. Brunecky R, Alahuhta M, Xu Q, Donohoe BS, Crowley MF, et al. (2013) Revealing nature's cellulase diversity: the digestion mechanism of Caldicellulosiruptor bescii CelA. Science 342: 1513-1516.
18. Mellitzer A, Glieder A, Weis R, Reisinger C, Flicker K (2012) Sensitive high-throughput screening for the detection of reducing sugars. Biotechnology Journal 7: 155-162.
19. Bien TLT, Tsuji S, Tanaka K, Takenaka S, Yoshida K-i (2014) Secretion of heterologous thermostable cellulases in Bacillus subtilis. The Journal of general and applied microbiology 60: 175-182.
20. Borchert TV, Nagarajan V (1991) Effect of signal sequence alterations on export of levansucrase in Bacillus subtilis. Journal of Bacteriology 173: 276-282.
21. Sakakibara Y, Tsutsumi K, Nakamura K, Yamane K (1993) Structural requirements of Bacillus subtilis alpha-amylase signal peptide for efficient processing: in vivo pulse-chase experiments with mutant signal peptides. Journal of Bacteriology 175: 4203-4212.

Chapter 6 Ethanol Production in *G. thermoglucosidasius* by Consolidated Bioprocessing

6.1 Introduction

Bioethanol is clean and renewable, which is the most-established biofuel. Its high cost still makes it less competitive, mainly because of large quantity of enzymes for saccharification. To eliminate the need of these enzymes in bioethanol production, a highly compact design of consolidated bioprocessing (CBP) is employed. It combines biomass degradation and sugar fermentation in a single step without enzyme addition, which is considered as a promising strategy for cost-effective ethanol production [1]. The key challenge of CBP is to develop a recombinant microorganism that is capable of producing enzymes for hydrolysis and fermenting the hydrolysate into ethanol.

Thermophilic *G. thermoglucosidasius* is a promising ethanol producer, because of advantages of high temperature growth. It is also a potential host for thermostable protein secretion, because it has high capability of exporting native proteins. By integrating these advantages, *G. thermoglucosidasius* can be developed for consolidated bioprocessing. In Chapter 3, ethanol production in *G. thermoglucosidasius* 95A1 was improved through metabolic engineering and evolutionary engineering. Several strains were achieved with high ethanol titer. In Chapter 5, *G. thermoglucosidasius* 95A1 was engineered for enhanced heterologous protein secretion by screening different signal peptides. As a result, optimal signal peptides for efficient secretion were found for the thermostable secretion targets of α-amylase and cellulases, respectively. In this chapter, work of Chapter 3 and work of Chapter 5 were combined to develop a consolidated bioprocessing for ethanol production from starch or cellulose. The CBP-

enabling strains were constructed by transforming amylase or cellulase plasmids with relevant optimal signal peptides into evolved strains from Chapter 3. As a result, the evolved strain with α-amylase secretion had efficient ethanol production directly from soluble starch; however, the evolved strains with different cellulase secretion produced little ethanol form cellulose. This chapter still needs more work to improve the strain for ethanol production by consolidated bioprocessing of cellulose.

6.2 Results

6.2.1 Construction of CBP-enabling strains for ethanol production

Starch is the important substrate for bioethanol production, which is an abundant renewable carbon source in nature. To develop a microorganism for consolidated bioprocessing starch for ethanol production, it requires the capability of producing key enzymes for starch hydrolysis, such as amylase, and also the ability of efficient fermentation for ethanol. In Chapter 3, engineered strains of *G. thermoglucosidasius* 95A1 were obtained with enhanced ethanol production from glucose and cellobiose. And in Chapter 5, a thermostable α-amylase was heterologously secreted in this strain and an efficient signal peptide was screened to direct the secretion at a high activity. Thus, a CBP-enabling strain of *G. thermoglucosidasius* was constructed by transforming the plasmid harboring truncated *amyE* with the optimal signal peptide (SP21-*amyE*) to the evolved strain for ethanol production (JZ03d2).

Cellulose is the most abundant carbohydrate polymer in nature, which is considered as the ideal substrate for bioethanol production. It requires different cellulase for biodegrading before being utilized for biofuel production. In Chapter 5, different thermostable cellulase was secreted in *G. thermoglucosidasius* and optimized by screening different SPs. Similarly, a CBP-enabling

strain was firstly constructed by transforming cellulase secretion plasmids of SP4-*celA* SP18-GH9/*celA* and SP16- GH48/*celA* separately to the evolved strain for ethanol production.

6.2.2 Direct ethanol production from starch

Wild type of *G. thermoglucosidasius* 95A1 can utilized starch [2], which is also verified in the results of Chapter 5 (Figure 5.2). Amylase secretion plasmid was transformed to fasten the consuming of starch and increase the ethanol productivity. The fermentation of strains was carried out in sealed conical tubes using the same conditions as those for ethanol production in Chapter 3. The only difference was the carbon source was replaced by 1% soluble starch. The *ldh* knockout strain (JZ01) was also tested as a comparison with the evolved strain. As seen in Figure 6.1, the secretion of amylase increased the growth rate and ethanol productivity in both of the strains. The production of the evolved strain with SP21-*amyE* reached 3.88±0.04 g/L ethanol after 48h, while the one without secreting amylase produced 2.35±0.05 g/L ethanol at the same time point. As same as the results in Chapter 3, evolved strains performed much better than the *ldh* knockout strains on ethanol production by using starch as the sole carbon source.

6.2.3 Direct ethanol production from cellulose

Phosphoric acid swollen cellulose (PASC) was used as the carbon source for fermentation. It is prepared from Avicel by phosphoric acid treatment. The cellulose swelling process changed its physical properties. Unlike Avicel, it is amorphous, and forms relatively viscous, cloudy, homogeneous solution in water. It is a model cellulose substrate used in studies about cellulase and cellulosic biofuel production. PASC was easier to be degraded than Avicel (Figure 6.2) because of loosen structure. From the screening results of cellulose in Chapter 5, *G. thermoglucosidasius* 95A1 didn't exhibit obvious cellulose activity. It was essential to have a small amount of glucose in the medium when the fermentation started. Therefore, the substrate

was hydrolyzed by GH9/CelA and GH48/CelA at 75°C for 24 h (Figure 6.3a). The enzyme for hydrolysis was inactivated or not before fermentation. The production of ethanol from PASC was carried out anaerobically in serum bottles using evolved strains with SP4-CelA, SP18-GH9/CelA and SP16-GH48/CelA secretion plasmids, respectively. Strains with SP18-GH9/CelA and SP16-GH48/CelA were inoculated together as a volume ratio of 1:2.

As shown in Figure 6.3b, the ethanol produced quickly in the first 12 h of fermentation by utilizing accumulated hydrolysis products. Afterwards, the ethanol production rate slowed down. The ones with hydrolysis enzyme keep producing ethanol in 72 h, although the productivity was very low. However, the others without active hydrolysis enzyme only produced very little before 48 h. Strains carrying cellulase plasmids performed better in ethanol production, but the difference was tiny. The reason for inefficient production should be that the cellulase was not active enough during fermentation. The optimal temperature of CelA was 75°C but the optimal temperature for ethanol production was 60°C.

6.3 Discussion and future directions

CBP was studied and developed a lot for ethanol production, because its efficient and economical process [1,3]. Much effort have been made to develop mesophilic microorganisms for CBP, such as *E. coli* [4], *S. cerevisiae* [5], *Z. mobilis* [6]. Thermophiles are attracted more interests for CBP developing because of benefits from high temperature growth. *G. thermoglucosidasius* has been reported for ethanol production [7,8] or enzyme expression and characterization [9,10], but not for CBP. Here, *G. thermoglucosidasius* was successfully engineered for improved ethanol production from starch by CBP for the first time. The promising result also confirmed the results of previous chapters: the efficient secretion of α-

amylase in 95A1 (Chapter 5) and improved ethanol production of the evolved strain (Chapter 3) on different substrates.

It was also the first trial for *G. thermoglucosidasius* for ethanol production from cellulose. The initial result showed that more work was required to increase the cellulolytic ability. The possible reasons were as follows: difference between the optimal temperatures, pH of CelA (75°C, pH 6.0) and ethanol production (60°C, pH 7.0); difference of expression systems from different species. Different fermentation conditions can be optimized, such as temperature, pH and medium. Besides, since the wild type has little activity on cellulose, higher initial OD_{600} can be tried for fermentation. When use the mixture of GH9/CelA and GH48/CelA, the composition can be also optimized. Although the inoculation was based on the optimal ratio from synergy assay, the two strains may have different growth rate.

Apart from optimizing conditions, secretion efficiency of enzymes can be improved. In *B. subtilis*, the main barrier to the production of heterologous protein secretion is the quality control proteases, especially the serine proteases (WprA, HtrA and HtrB) [11,12,13]. According to homologs analysis, several serine proteases were found in *G. thermoglucosidasius* 95A1. The deletion of these genes is still undergoing.

The key challenge of making this strain being CBP-enabling is to endow it with the ability of degrading cellulose efficiently. Here, CelA from *C. bescii* was tried. Although it is the most highly secreted cellulase in *C. bescii*, it looks not effective enough for *G. thermoglucosidasius*. More cellulose from thermophilic bacteria can be tried, such as *T. maritima* and *C. thermocellum*. Besides, many cellulolytic microorganisms in natural secrete cellulosome (multicellulase complex), associated with the cell surface, for efficient breakdown of cellulose. This may be another way to improve the ability of cellulose degradation in this thermophile.

Here, ethanol was chosen as the product because it is an important biofuel and its production benefits a lot from high temperature fermentation. From the experiment design, it is more convenient for strain construction. The strain constructed for *R*-BDO production (in Chapter 4) can also be used for CBP by using starch or cellulose. The constructs for hydrolase secretion can be integrated into genome, which can avoid the potential problem of plasmids compatibility.

6.4 Conclusion

In conclusion, *G. thermoglucosidasius* 95A1 was engineered as a CBP-enabling host for the first time producing ethanol from starch and cellulose, respectively. Promising result was obtained by utilizing starch as substrate through secreting a thermostable amylase fused with an efficient signal peptide in engineered strain for ethanol production. More work is needed for cellulosic biofuel production by CBP, mainly developing more efficient cellulose degradation system in this strain. All in all, this chapter established this thermophile as a cell factory for biofuel and chemical production.

6.5 Material and methods

6.5.1 Ethanol production from starch

Fermentation was carried out in15 mL sealed conical tubes. The conditions were the same as fermentation for ethanol production from glucose (described in 3.5.4), but 1% soluble starch (Fisher Scientific) was used as the sole carbon source. The product was analyzed by HPLC.

6.5.2 Ethanol production from PASC

PASC was freshly prepared (described in 2.3.3) and autoclaved. It was hydrolyzed by secreted enzyme of SP18-GH9/CelA and SP16-GH48/CelA before using as the substrate. Two strains were cultured in TGP at 55°C, 250 rpm until OD_{600} reached around 4. Then 50 mL

supernatant of SP18-GH9/CelA and 100 mL supernatant of SP16-GH48/CelA were concentrated by using Amicon Ultra Centrifugal Filters and then filtered through 0.22 μm filter. The hydrolysis was performed in CelA buffer (50 mM sodium-citrate, 150 mM NaCl, pH 6.0) with freshly prepared 2% PASC and concentrated supernatant of culture. The reaction was incubated for 24 h at 75°C, 150 rpm. After reaction, the enzymes were inactivated (by boiling) or not before adding to the medium as indicated.

Seed culture for ethanol production was prepared in 500 mL flasks with 100 mL TGP. The culture was inoculated with a single colony from a fresh plate and grown aerobically at 60°C, 250 rpm for 12 hour. Cells harvested were washed with PBS (pH 7.0) and resuspended into TMLM medium with 1% hydrolyzed PASC and 0.5% yeast extract to initial OD_{600} of 0.5. Fermentation was carried out in 30 mL sealed serum bottle containing 15 mL medium at 60 °C, 250 rpm. 0.5 mL of sample was taken at time intervals as indicated and analyzed by HPLC.

6.6 Table

Table 6.1 Plasmids and strains used in this study

Plasmid or strain	Characteristics	Source or reference
Plasmids		
SP21-*amyE*	DNA fragment of SP21 fused with truncated *amyE* in pJZ57	The 5th Chapter
SP4-*celA*	DNA fragment of SP4 fused with truncated *celA* in pJZ83	The 5th Chapter
SP18-*GH9/celA*	DNA fragment of SP18 fused with truncated *GH9/celA* in pJZ85	The 5th Chapter
SP16- *GH48/celA*	DNA fragment of SP16 fused with truncated *GH48/celA* in pJZ87	The 5th Chapter
Strains		
JZ01	Δ*ldh* variant of *G. thermoglucosidasius* 95A1	The 3rd chapter
JZ03d2	Evolved *Δldh Δpfl* variant of *G. thermoglucosidasius* 95A1	The 3rd chapter

6.7 Figures

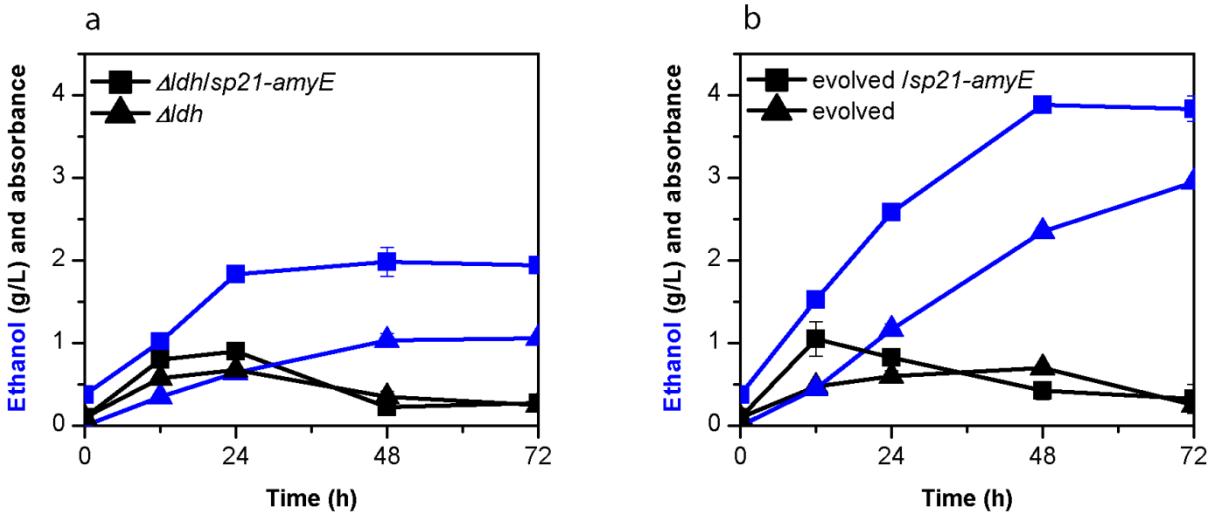

Figure 6.1 Direct ethanol production from starch in engineered *G. thermoglucosidasius* 95A1. Cells were cultured in TMLM with 1% soluble starch at 60°C. (a) Ethanol production and cell growth in Δ*ldh* (JZ01) and Δ*ldh/sp21-amyE*. (b) Ethanol production in evolved strain (JZ03d2) evolved strain with *sp21-amyE*. The figure shows the means and standard deviations obtained from three experiments.

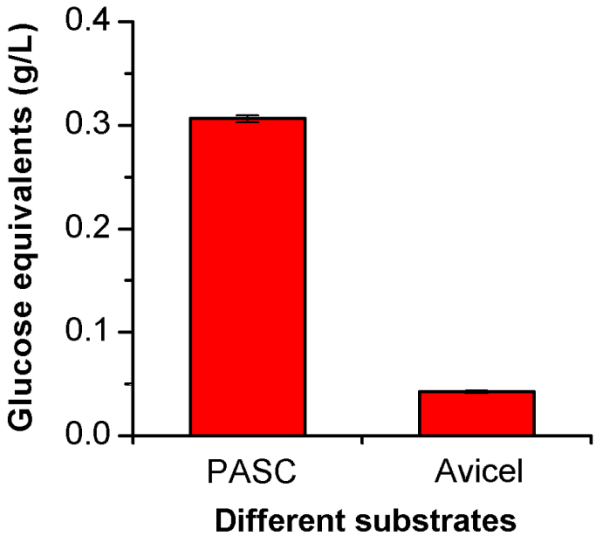

Figure 6.2 CelA activity on different substrates. Enzyme reaction was performed at 75°C for 24h, using 1% PASC or Avicel as the substrate. Enzyme activity was determined by pHBAH assay. The figure shows the means and standard deviations obtained from three experiments.

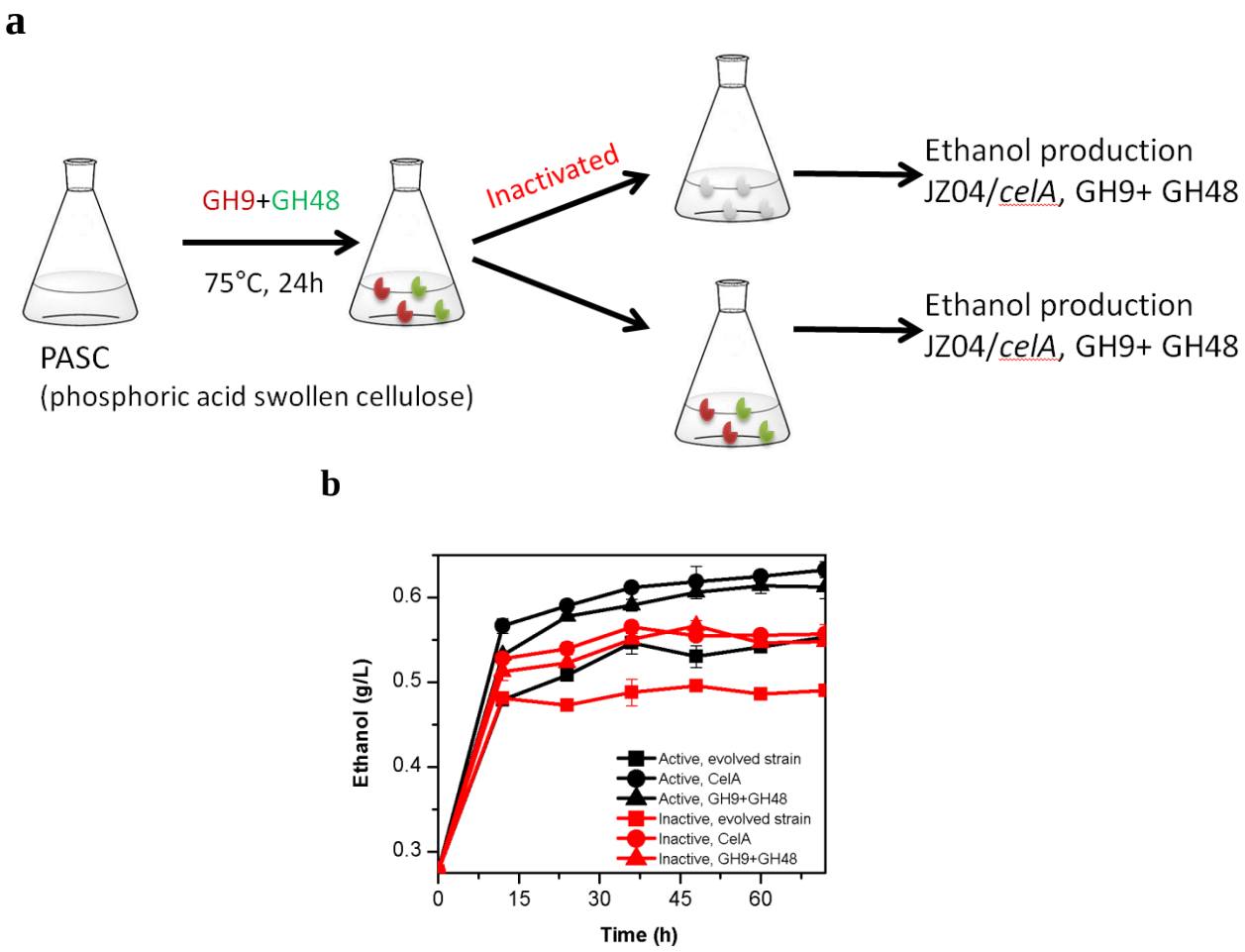

Figure 6.3 Direct ethanol production from cellulose in engineered *G. thermoglucosidasius* 95A1. a, schematic overview of ethanol production from cellulose. 1% PASC was used as the substrate, which was hydrolyzed by GH9/CelA and GH48/CelA at 75°C for 24h. Enzyme was inactivated or not before fermentation started. b, fermentation of cellulose for ethanol production. Cells were cultured in TMLM with 0.5% yeast extract and 1% PASC, at 60°C. Red curves indicated the enzyme added before fermentation was inactivated, and black curves indicated the enzyme wasn't treated. The figure shows the means and standard deviations obtained from three experiments.

6.8 References

1. Olson DG, McBride JE, Shaw AJ, Lynd LR (2012) Recent progress in consolidated bioprocessing. Current Opinion in Biotechnology 23: 396-405.
2. Zeigler DR (2001) The genus Geobacillus. Introduction and strain catalog. Catalog of strains 3.
3. Lynd LR, Van Zyl WH, McBride JE, Laser M (2005) Consolidated bioprocessing of cellulosic biomass: an update. Current Opinion in Biotechnology 16: 577-583.
4. Bokinsky G, Peralta-Yahya PP, George A, Holmes BM, Steen EJ, et al. (2011) Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered Escherichia coli. Proceedings of the National Academy of Sciences 108: 19949-19954.
5. Wen F, Sun J, Zhao H (2010) Yeast surface display of trifunctional minicellulosomes for simultaneous saccharification and fermentation of cellulose to ethanol. Applied and Environmental Microbiology 76: 1251-1260.
6. Wang G-J, Wang Z-S, Zhang Y-W, Zhang Y-Z (2012) Cloning and expression of amyE gene from Bacillus subtilis in Zymomonas mobilis and direct production of ethanol from soluble starch. Biotechnology and bioprocess engineering 17: 780-786.
7. Cripps RE, Eley K, Leak DJ, Rudd B, Taylor M, et al. (2009) Metabolic engineering of Geobacillus thermoglucosidasius for high yield ethanol production. Metabolic Engineering 11: 398-408.
8. Van Zyl LJ, Taylor MP, Eley K, Tuffin M, Cowan DA (2013) Engineering pyruvate decarboxylase-mediated ethanol production in the thermophilic host Geobacillus thermoglucosidasius. Applied Microbiology and Biotechnology 98: 1247-1259.
9. Bartosiak-Jentys J, Hussein AH, Lewis CJ, Leak DJ (2013) Modular system for assessment of glycosyl hydrolase secretion in Geobacillus thermoglucosidasius. Microbiology 159: 1267-1275.
10. Fong SS, Pogrebnyakov I, Jendresen CB, Nielsen AT (2017) Genetic toolbox for controlled expression of functional proteins in Geobacillus spp. PloS one 12: e0171313.
11. Harwood CR, Cranenburgh R (2008) Bacillus protein secretion: an unfolding story. Trends in Microbiology 16: 73-79.
12. Darmon E, Noone D, Masson A, Bron S, Kuipers OP, et al. (2002) A novel class of heat and secretion stress-responsive genes is controlled by the autoregulated CssRS two-component system of Bacillus subtilis. Journal of Bacteriology 184: 5661-5671.
13. Stephenson K, Harwood CR (1998) Influence of a cell-wall-associated protease on production of α-amylase by Bacillus subtilis. Applied and Environmental Microbiology 64: 2875-2881.